

From: O'Hara, Eileen  
Sent: Wednesday, September 18, 2002 5:39 PM  
To: STIC-Biotech/ChemLib  
Subject: sequence search request for 09/746,375

## SEARCH REQUEST FORM-SEQUENCE

Examiner: Eileen O'Hara (Room 10E18) Art Unit 1646

Serial Number: **09/746,375**  
**Novel Cytokine ZCYTO18**  
**Presnell et al**

Date: 09/18/02 Phone: 308-3312 CM1 10A01 mailbox 10C01

Please search the following:

SEQ ID NO: 1

also, 14 contiguous nucleotides of SEQ ID NO: 1

in commercial and interference databases.

Please put results on paper.

Thank you!

Eileen

**Eileen B. O'Hara**  
**CM1 10A01 mailbox 10C01**  
**703-308-3312**  
**U.S. Patent and Trademark Office**  
**Art Unit 1646**  
**eileen.o'hara@uspto.gov**

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

RECEIVED  
SEP 19 2002  
STIC/Biotech (STIC)

10019

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 9/19/02  
Date Completed: 10/3/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: 2  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: 04  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2002, 00:43:59 ; Search time 217.46 Seconds  
(without alignments)  
8811.168 Million cell updates/sec

Title: US-09-746-375-1

Perfect score: 1116

Sequence: 1 tcgagtagtaattgtctgca.....ataaacacttgatataocta 1116

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 14

Total number of hits satisfying chosen parameters: 19226

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_032802.\*

1: /SIDS1/gcgdata/genseq/genseq-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/genseq/genseq-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/genseq/genseq-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/genseq/genseq-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/genseq/genseq-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/genseq/genseq-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/genseq/genseq-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/genseq/genseq-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/genseq/genseq-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/genseq/genseq-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/genseq/genseq-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/genseq/genseq-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/genseq/genseq-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/genseq/genseq-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/genseq/genseq-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/genseq/genseq-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/genseq/genseq-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/genseq/genseq-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/genseq/genseq-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/genseq/genseq-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/genseq/genseq-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/genseq/genseq-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/genseq/genseq-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/genseq/genseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1116	100.0	1116 22 AAD09719	Human cytokine, ZC
2	1116	100.0	1116 22 AAF83741	Human IL-TIF polyp
3	1116	100.0	1116 22 AAF81773	Human GIL-19/AE289
4	1114	99.8	1152 22 AAF92134	Human PRQ10096 cDN
5	1114	99.8	1152 22 AAC87053	Nucleotide sequenc
6	1090	97.7	1132 22 AAC84310	Human EXCS encodin
7	934	83.7	1139 22 AAF28841	Human interleukin-
8	639	57.3	690 22 AAF28839	Human T cell induc
9	639	57.3	690 22 AAS14875	Human cDNA encodin

10	207	18.5	4796	21	AAA28840	Human T cell induc
11	207	18.5	4797	22	AAS14876	Human partial geno
12	35	3.1	778	22	AAD09746	Mouse ZCYTO18 cDNA
13	35	3.1	1119	21	AAA28815	Murine T cell indu
14	35	3.1	1119	22	AAS14858	Mouse cDNA encodin
15	35	3.1	1166	21	AAC81774	Murine GIL-19 prot
16	31	2.8	7444	21	AAA28816	Murine T cell indu
17	31	2.8	7445	22	AAS14859	Mouse partial geno
18	28	2.5	1111	21	AAA28817	Murine T cell indu
19	28	2.5	1111	22	AAS14860	Mouse cDNA for T c
20	28	2.5	5935	21	AAA28818	Murine T cell indu
21	28	2.5	5935	22	AAS14878	Mouse partial geno
22	27	2.4	36	22	AAD09730	Human ZCYTO18 cDNA
23	27	2.4	36	22	AAF83753	Human IL-TIF cDNA
24	24	2.2	24	22	AAD09721	PCR primer, ZC2584
25	24	2.2	24	22	AAD09722	PCR primer, ZC2584
26	24	2.2	33	22	AAD09731	Human ZCYTO18 cDNA
27	24	2.2	33	22	AAF83754	Human IL-TIF cDNA
28	24	2.2	5421	24	ABL32132	Human immune syste
29	23	2.1	23	22	AAD09724	ZC25839 PCR primer
30	23	2.1	37	22	AAD09734	Human ZCYTO18 cDNA
31	22	2.0	22	21	AAA28831	Primer for human T
32	22	2.0	22	22	AAS14867	Human T cell deriv
33	22	2.0	39	22	AAD09733	Human ZCYTO18 cDNA
34	22	2.0	2640	22	AAH76178	A. thaliana SGT101
35	22	2.0	2671	22	AAH76176	A. thaliana SGT101
36	21	1.9	21	21	AAA28833	5' RACE primer for
37	21	1.9	21	21	AAA28834	5' RACE primer for
38	21	1.9	21	22	AAS14869	Human T cell deriv
39	21	1.9	21	22	AAS14870	Human T cell deriv
40	21	1.9	706	22	AAH57394	Human skeletal mus
41	21	1.9	3126	21	AAA70143	Plasmodium falcipa
42	21	1.9	15768	24	ABL33171	Human immune syste
43	21	1.9	16217	24	ABL32625	Human immune syste
44	20	1.8	473	22	ABA43453	Human breast cell
45	20	1.8	473	22	ABA53906	Human foetal liver

#### ALIGNMENTS

RESULT 1  
AAD09719  
ID AAD09719 standard; cDNA; 1116 BP.  
XX  
AC AAD09719;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Human cytokine, ZCYTO18 cDNA.  
XX  
KW Human; cytostatic; cytokine; ZCYTO18 protein; genetic abnormality;  
cancer; inflammation; gene therapy; chromosome 12; ss.  
XX  
OS Homo sapiens.  
XX  
Key Location/Qualifiers  
FT CDS 21..560  
FT /\*tag= a  
FT /\*product= "Human ZCYTO18 protein #1"  
FT sig\_peptide 21..119  
FT /\*tag= b  
FT mat\_peptide 120..557  
FT /\*tag= c  
FT /\*product= "Human mature ZCYTO18 protein #1"  
FT CDS 57..560  
FT /\*tag= d  
FT /\*product= "Human ZCYTO18 protein #2"  
FT sig\_peptide 57..119  
FT /\*tag= e  
FT mat\_peptide 120..557  
FT /\*tag= f  
FT /\*product= "Human mature ZCYTO18 protein #2"

XX WO200146422-A1.  
PN 28-JUN-2001.  
PS 22-DEC-2000; 2000WO-US35308.  
PP 23-DEC-1999; 99US-0471767.  
PR 01-DEC-2000; 2000US-0250841.  
XX (ZYMO ) ZYMOGENETICS INC.  
XX Presnell SR, Kindsvogel W;  
PI WPI; 2001-408648/43.  
DR P-PSDB; AAE04538, AAE04539.  
XX Novel human cytokine polypeptide, ZCYTO18, useful for treating cancer  
XX Claim 2; Page 140-142; 167pp; English.  
XX The patent discloses novel human cytokine, ZCYTO18 protein and its  
CC corresponding DNA. ZCYTO18 protein induces proliferation of cells  
CC expressing zcytor11, a receptor for ZCYTO18 or induces cytotoxicity  
CC in K562 cells. ZCYTO18 DNA is useful for detecting a genetic  
CC abnormality in a patient. ZCYTO18 DNA and its antibodies are useful  
CC for detecting cancer and inflammation. ZCYTO18 protein is useful for  
CC killing cancer cells. It is useful for increasing platelets in a  
CC patient or injured tissue. It is also used in gene therapy.  
CC The present sequence is a cDNA encoding novel human cytokine, ZCYTO18  
CC protein. ZCYTO18 DNA is located at the 12q15 region of chromosome 12.  
XX Sequence 1116 BP; 336 A; 229 C; 222 G; 329 T; 0 other;  
O;  
Query Match 100.0%; Score 1116; DB 22; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
O;  
Y 1 tcgaattagaattgtctgcaatggccgctccttcctgcagaaatctgtgagctcttccctatgg 60  
D 1 tcgaattagaattgtctgcaatggccgctccttcctgcagaaatctgtgagctcttccctatgg 60  
Y 61 ggaacctggccacagctgcctcctctctctctctctctctctctctctctctctctctct 120  
D 61 ggaacctggccacagctgcctcctctctctctctctctctctctctctctctctctctct 120  
Y 121 cgcacctacgctccacagctgcagctgcagctgcagctgcagctgcagctgcagctgcagct 180  
D 121 cgcacctacgctccacagctgcagctgcagctgcagctgcagctgcagctgcagctgcagct 180  
Y 181 accgcaccttcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 240  
D 181 accgcaccttcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 240  
Y 241 tcaattggggagaaactgttccacgagctcagctcagctcagctcagctcagctcagctcagc 300  
D 241 tcaattggggagaaactgttccacgagctcagctcagctcagctcagctcagctcagctcagc 300  
Y 301 aggtgctgaactccacctggaagctgctgttccctcaatctgataggttccagcctt 360  
D 301 aggtgctgaactccacctggaagctgctgttccctcaatctgataggttccagcctt 360  
Y 361 atatgcaggagggtgtgtcctcctgctgctgctgctgctgctgctgctgctgctgctgct 420  
D 361 atatgcaggagggtgtgtcctcctgctgctgctgctgctgctgctgctgctgctgctgct 420  
Y 421 ttgaaggtgatgacctgcctatccagagaaatgtgcaaaagctgaaagacacagtgaaaaa 480  
D 421 ttgaaggtgatgacctgcctatccagagaaatgtgcaaaagctgaaagacacagtgaaaaa 480  
Y 481 agcttgagagagctgagagatcaaaagcaattggagaactggatttgcgttatgtctc 540  
D 481 agcttgagagagctgagagatcaaaagcaattggagaactggatttgcgttatgtctc 540

Db 481 agcttgagagagctgagagatcaaaagcaattggagaactggatttgcgttatgtctc 540  
Qy 541 tgagaaatgctgctcatttgaccagagcaaaagctgagaaatgaataactaacccctttcc 600  
Db 541 tgagaaatgctgctcatttgaccagagcaaaagctgagaaatgaataactaacccctttcc 600  
Qy 601 ctgctagaaataaacaattagatgcccccaagcgaatttttttaaccacaaaggagatggg 660  
Db 601 ctgctagaaataaacaattagatgcccccaagcgaatttttttaaccacaaaggagatggg 660  
Qy 661 aagccaaatcccatcatgatgggtggattccaaatgaacccctgcgttagttacaaagga 720  
Db 661 aagccaaatcccatcatgatgggtggattccaaatgaacccctgcgttagttacaaagga 720  
Qy 721 aaccaatgcccactttgtttataagaccagagtagaactttcttaagcatagatatattat 780  
Db 721 aaccaatgcccactttgtttataagaccagagtagaactttcttaagcatagatatattat 780  
Qy 781 tgataacatttcattgtaactggtgtctctatcacagagaaacaaatttttttaaatat 840  
Db 781 tgataacatttcattgtaactggtgtctctatcacagagaaacaaatttttttaaatat 840  
Qy 841 ttgctttttccataaaaaaagattacttccattcccttttaggggaaaaaacccctaaata 900  
Db 841 ttgctttttccataaaaaaagattacttccattcccttttaggggaaaaaacccctaaata 900  
Qy 901 gcttcattgctccataacacagcttctattataatgataatgattattattattataaga 960  
Db 901 gcttcattgctccataacacagcttctattataatgataatgattattattattataaga 960  
Qy 961 ctgctttttattatcatcttttattataatgataatgattattattattattattattat 1020  
Db 961 ctgctttttattatcatcttttattataatgataatgattattattattattattattat 1020  
Qy 1021 ttgcttacttgagtgtaaggctaaatttattattattattattattattattattattattat 1080  
Db 1021 ttgcttacttgagtgtaaggctaaatttattattattattattattattattattattattat 1080  
Qy 1081 tgttatttgacctcaataaacaacttgatattccta 1116  
Db 1081 tgttatttgacctcaataaacaacttgatattccta 1116  
RESULT 2  
AAAF3741  
ID AAF3741 standard; cDNA; 1116 BP.  
XX  
AC AAF3741;  
XX  
DT 23-JUL-2001 (first entry)  
XX  
DE Human IL-TIF polypeptide encoding cDNA.  
XX  
XX  
KW Cytokine receptor; zcytor16; IL-TIF; antinflammatory; cytostatic;  
KW antirheumatic; antiarthritic; antiasthmatic; antiatherosclerotic;  
KW immunosuppressive; chromosome 6q24.1-25.2; human; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 21..560  
FT /\*tag= a  
FT /product= "IL-TIF"  
XX  
XX WO200140467-A1.  
PN  
PD 07-JUN-2001.  
XX  
XX 01-DEC-2000; 2000WO-US32703.  
XX  
XX 03-DEC-1999; 99US-0169049.  
PR 13-SEP-2000; 2000US-0232219.  
PR 31-OCT-2000; 2000US-0244610.  
PR





XX Human GIL-19 protein that shows a high degree of homology to IL  
PT (interleukin)-10, useful in upregulation of humoral immune responses,  
PT as an antiinflammatory agent and as a modulator of immune responses  
PT associated with injury -

XX Claim 1: Page 59; 60pp; English.

XX The present invention provides the protein and coding sequences for the  
CC novel human GIL-19/AE289 protein. The protein shows homology to  
CC interleukin-10 (IL-10) and is assumed to be a cytokine. It can be used  
CC in the regulation of cell proliferation and differentiation,  
CC haematopoiesis, immune stimulation or suppression, tissue growth and  
CC tumour inhibition. In addition, it also has uses in the treatment of  
CC inflammation and in nutrition.

XX Sequence 1177 BP; 362 A; 245 C; 232 G; 338 T; 0 other;

Query Match 100.0%; Score 1116; DB 21; Length 1177;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcgagtttagaattgtctgcgaatggccgcgcctgcgaagaatctgtgagctcttccattgg 60  
DB 45 tcgagtttagaattgtctgcgaatggccgcgcctgcgaagaatctgtgagctcttccattgg 104

QY 61 ggcacctggccaccagctgcctctctcttggccctcttggtacagggagagcagctg 120  
DB 105 ggcacctggccaccagctgcctctctcttggccctcttggtacagggagagcagctg 164

QY 121 cgcccatcagctcccaatgcaggcttgacagttcacaattccagcagcctatatacca 180  
DB 165 cgcccatcagctcccaatgcaggcttgacagttcacaattccagcagcctatatacca 224

QY 181 accgcaccttcattcgttggaagagcagcttggtgatacaacacagacgttcgtc 240  
DB 225 accgcaccttcattcgttggaagagcagcttggtgatacaacacagacgttcgtc 284

QY 241 tcattggggagaaactgtccacagagtcagatgagtgagcgtctgtctatgatgaagc 300  
DB 285 tcattggggagaaactgtccacagagtcagatgagtgagcgtctgtctatgatgaagc 344

QY 301 aggtgctgaacttcaccccttgagaagaatgctgttccctcattcgtatgagttccagctc 360  
DB 345 aggtgctgaacttcaccccttgagaagaatgctgttccctcattcgtatgagttccagctc 404

QY 361 atatgcaggaggtggtgcttctcctggccaggtcagcaacagcgttaagcacatgctata 420  
DB 405 atatgcaggaggtggtgcttctcctggccaggtcagcaacagcgttaagcacatgctata 464

QY 421 ttgaaggtgatgacctgcataatcccgaggaatgtgcgaagaagctgaaggacacagtgaaaa 480  
DB 465 ttgaaggtgatgacctgcataatcccgaggaatgtgcgaagaagctgaaggacacagtgaaaa 524

QY 481 agcttgagagagtgagagatcaaaagcaatggagaaactggattgtgtttatgtctc 540  
DB 525 agcttgagagagtgagagatcaaaagcaatggagaaactggattgtgtttatgtctc 584

QY 541 tgagaaatgctgcatttgacacagagcaagctgaaatgaataactaacccctttcc 600  
DB 585 tgagaaatgctgcatttgacacagagcaagctgaaatgaataactaacccctttcc 644

QY 601 ctgctagaaaaaacaatttagatgcccccaagcgaatttttttaacaaaagaagatggg 660  
DB 645 ctgctagaaaaaacaatttagatgcccccaagcgaatttttttaacaaaagaagatggg 704

QY 661 aagccaaactccatcatgatgggtggattccaaaagaaccctcggttagttacaaga 720  
DB 705 aagccaaactccatcatgatgggtggattccaaaagaaccctcggttagttacaaga 764

QY 721 aaccatgcacctttgtttataacagacagaggttagacctttcctaagcatagatatattat 780

DB 765 aaccaatgccacttttgtttataaagaccaggaaggtagaactttcttaagcatagatttat 824  
QY 781 tgatacaattctattgtaactgtgttctctatcacagagaaacaatttttttaataaa 840  
DB 825 tgatacaattctattgtaactgtgttctctatcacagagaaacaatttttttaataaa 884

QY 841 ttgtctttttccataaaaaagattactcttccattcccttaggggaaaaaaccccttaata 900  
DB 885 ttgtctttttccataaaaaagattactcttccattcccttaggggaaaaaaccccttaata 944

QY 901 gctcatgtttccataatcagtgactttattattataaatgtattattattattataaga 960  
DB 945 gctcatgtttccataatcagtgactttattattataaatgtattattattattataaga 1004

QY 961 ctgcattttattatcatcattttattataatgatgattattattatagaacaatcattcgata 1020  
DB 1005 ctgcattttattatcatcattttattataatgatgattattattatagaacaatcattcgata 1064

QY 1021 ttgctacttgagtgtaagctagatttgatatttatgataattatgacaataattatagagctataaca 1080  
DB 1065 ttgctacttgagtgtaagctagatttgatatttatgataattatgacaataattatagagctataaca 1124

QY 1081 tgtttatttgacctcaataaaacaccttggtatcccta 1160  
DB 1125 tgtttatttgacctcaataaaacaccttggtatcccta 1160

RESULT 4  
AAF921134  
ID AAF92134 standard; cDNA; 1152 BP.

XX AC AAF92134;  
XX DT 15-MAY-2001 (first entry)  
XX DE Human PRO10096 cDNA.  
XX KW Human; PRO protein; mapping; ss.  
XX OS Homo sapiens.  
XX PN WO200116318-A2.  
XX PD 08-MAR-2001.  
XX PF 24-AUG-2000; 2000WO-US23328.  
XX PR 01-SEP-1999; 99WO-US20111.  
XX PR 15-SEP-1999; 99WO-US21090.  
XX PR 07-DEC-1999; 99US-0169495.  
XX PR 09-DEC-1999; 99US-0170262.  
XX PR 11-JAN-2000; 2000US-0175481.  
XX PR 18-FEB-2000; 2000WO-US04341.  
XX PR 18-FEB-2000; 2000WO-US04342.  
XX PR 22-FEB-2000; 2000WO-US04414.  
XX PR 01-MAR-2000; 2000WO-US05601.  
XX PR 03-MAR-2000; 2000US-0187202.  
XX PR 25-APR-2000; 2000US-0199397.  
XX PR 22-MAY-2000; 2000WO-US14042.  
XX PR 05-JUN-2000; 2000US-0209832.  
XX PA (GETH ) GENENTECH INC.  
XX PI Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;  
XX WPI; 2001-183260/18.  
XX P-PSDB; AAB87602.  
XX PT Eighty four nucleic acids encoding PRO polypeptides, useful in  
PT molecular biology, including use as hybridization probes, and in  
PT chromosome and gene mapping.

PS Claim 2; Fig 153; 278pp; English.

XX The present sequence is the coding sequence for a human PRO polypeptide  
CC (secreted and transmembrane). The PRO protein, and PRO agonists, PRO  
CC antagonists or anti-PRO antibodies are useful for preparation of a  
CC medicament useful in the treatment of a condition which is responsive to  
CC the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO  
CC protein may also be employed as molecular weight markers for protein  
CC electrophoresis. The PRO coding sequence has applications in molecular  
CC biology, including use as hybridisation probes, and in chromosome and  
XX gene mapping.

SQ Sequence 1152 BP; 342 A; 244 C; 228 G; 338 T; 0 other;

Query Match 99.8%; Score 1114; DB 22; Length 1152;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcagagttagaattgtctgcaatggcgccctgcagaaatctgtgagctttcccttatgg 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
38 tcagagttagaattgtctgcaatggcgccctgcagaaatctgtgagctttcccttatgg 97  
QY 61 ggaacctgagccacagctgcctccctctcttggccctctgtgtacaggagagcagctg 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
98 ggaacctgagccacagctgcctccctctcttggccctctgtgtacaggagagcagctg 157  
QY 121 cggccatcagctcccaactgcagcttgacaagctccaactccagcagccctatacaca 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
158 cggccatcagctcccaactgcagcttgacaagctccaactccagcagccctatacaca 217  
QY 181 accgcaactctatgtctgctgaaggagctgtgctgtgatacaacacagacgttcgctc 240  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
218 accgcaactctatgtctgctgaaggagctgtgctgtgatacaacacagacgttcgctc 277  
QY 241 tcattggggagaaactgttccacgagtcagctatgagtgaggcgtgctatctgtatgaagc 300  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
278 tcattggggagaaactgttccacgagtcagctatgagtgaggcgtgctatctgtatgaagc 337  
QY 301 aggtgtgaacttcaccccttgaagagtgctgttccctcaactctgatggttccagcctt 360  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
338 aggtgtgaacttcaccccttgaagagtgctgttccctcaactctgatggttccagcctt 397  
QY 361 atatcaggaggtgtgctcctccctgcccagctcagcaacaggctaagcagcatgtcata 420  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
398 atatcaggaggtgtgctcctccctgcccagctcagcaacaggctaagcagcatgtcata 457  
QY 421 ttgaaggtgatgacctgcataatccagaggaatgtgcaaaagctgaaggacagtgaaaa 480  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
458 ttgaaggtgatgacctgcataatccagaggaatgtgcaaaagctgaaggacagtgaaaa 517  
QY 481 agcttggagagagtgagagatcaaaagcaatgggagaaactggttctgtttatgtctc 540  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
518 agcttggagagagtgagagatcaaaagcaatgggagaaactggttctgtttatgtctc 577  
QY 541 tgagaaatgcctgtcatttgaccagagcaaatgaaatgaatacctacccctttcc 600  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
578 tgagaaatgcctgtcatttgaccagagcaaatgaaatgaatacctacccctttcc 637  
QY 601 ctgctagaaataaactatagatgcccaagcgatatttttaaccacaaaggaagatggg 660  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
638 ctgctagaaataaactatagatgcccaagcgatatttttaaccacaaaggaagatggg 697  
QY 661 aagccaaactcctcatcatgatgggtggatcccaatgaacccctgcgttagttacaaagga 720  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
698 aagccaaactcctcatcatgatgggtggatcccaatgaacccctgcgttagttacaaagga 757  
QY 721 aaccaatgccactttgtttataagaccagaaggtagactttcttaagcatagatatattat 980  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
758 aaccaatgccactttgtttataagaccagaaggtagactttcttaagcatagatatattat 817  
QY 781 tgataaacatttctatgtactgtgtctctatacacagaaaaaacaaatttttttaataa 840  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 818 tgataacatttctgttaactgtgttctatatacagaaaaacaattattttttaataa 877  
QY 841 ttgtctttttccataaaaaagattacttttccattctttaggggaaaaaacccctaaata 900  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
878 ttgtctttttccataaaaaagattacttttccattctttaggggaaaaaacccctaaata 937  
QY 901 gttcatgtttccataataatcagtcactttattattataaaatgtattattattataaga 960  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
938 gttcatgtttccataataatcagtcactttattattataaaatgtattattattataaga 997  
QY 961 ctcgcatattttattatcatattttattataatgtgattttattatagaacacattcgata 1020  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
998 ctcgcatattttattatcatattttattataatgtgattttattatagaacacattcgata 1057  
QY 1021 ttgtctacttgagtgtaaggctaatatttgattttatgacaataattatagactataaca 1080  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1058 ttgtctacttgagtgtaaggctaatatttgattttatgacaataattatagactataaca 1117  
QY 1081 tgtttatttgacctcaataaacaacttggatatcc 1114  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1118 tgtttatttgacctcaataaacaacttggatatcc 1151  
RESULT 5  
AAC87053  
ID AAC87053 standard; cDNA; 1152 BP.  
XX  
AC AAC87053;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Nucleotide sequence of human polypeptide PRO10096.  
XX  
KW Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;  
KW PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1361; PRO1308;  
KW PRO118; PRO1272; PRO1419; PRO499; PRO170; PRO248; PRO353; PRO1318;  
KW PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO1411; PRO4356;  
KW PRO246; PRO265; PRO341; PRO10096; PRO6003; PRO6004; PRO350; PRO2630;  
KW PRO6309; cell death; genetic disorder; transgenic animal; gene therapy;  
KW ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 58..597  
FT sig\_peptide 58..156 /\*tag= a  
FT /\*tag= b  
XX  
PN WO200077037-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 22-MAY-2000; 2000WO-US14042.  
XX  
PR 15-JUN-1999; 99US-0139695.  
PR 20-JUL-1999; 99US-0145070.  
PR 26-JUL-1999; 99US-0145698.  
PR 17-AUG-1999; 99US-0149396.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28301.  
PR 02-DEC-1999; 99WO-US28565.  
PR 07-DEC-1999; 99US-0169495.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05841.  
PR

PR 20-MAR-2000: 2000WO-US07377.  
 PR 30-MAR-2000: 2000WO-US08439.  
 PR 15-MAY-2000: 2000WO-US13358.  
 PR 17-MAY-2000: 2000WO-US13705.

XX (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Botstein DA, Desnovers L, Eaton DL,  
 PI Ferrara N, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A,  
 PI Godowski PJ, Gurney AL, Kijavlin LJ, Mather JP, Napier MA, Pan J,  
 PI Pooni NF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM,  
 PI Wood WI, Zhang Z;

XX WPI: 2001-050091/06.  
 DR P-PSDB: AAB31210.

XX Isolated nucleic acid molecule encoding a PRO polypeptide which is a  
 PT transmembrane polypeptide is useful for gene therapy and identification  
 PT of related polypeptides -

XX Claim 2: Fig 63: 244pp: English.

CC The present sequence encodes a human secreted and transmembrane  
 CC polypeptide. The specification describes human polypeptides, designated  
 CC PRO196, PRO444, PRO183, PRO185, PRO210, PRO217, PRO242, PRO286,  
 CC PRO365, PRO1361, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO7170,  
 CC PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,  
 CC PRO337, PRO1411, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,  
 CC PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells  
 CC can be modulated with agents that bind to these polypeptides, resulting  
 CC in the death of the cells. The polynucleotides encoding of these  
 CC polypeptides are useful in the recombinant production of the  
 CC polypeptides, as a hybridisation probe to screen libraries to isolate  
 CC homologous sequences, or to map the gene. They may also be used for  
 CC analysing genetic disorders, and to produce transgenic animals which are  
 CC useful for the development and screening of therapeutically useful  
 CC reagents. The polynucleotides can also be used in gene therapy e.g. to  
 CC replace a defective gene.

XX Sequence 1152 BP: 342 A: 244 C: 228 G: 338 T: 0 other:

Query Match 99.8%; Score 1114; DB 22; Length 1152;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tcgaagtagaattgtctgcaatggccgctgcagaaatctgtgagctcttccctatgg 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 38 tcgaagtagaattgtctgcaatggccgctgcagaaatctgtgagctcttccctatgg 97  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 61 ggaccctggccaccagctgctctctcttctggccctcttggtacaggaggagcagctg 120  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 98 ggaccctggccaccagctgctctctcttctggccctcttggtacaggaggagcagctg 157  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 121 ggcctatcagctcccaactgcaggctgtgcaagtcacaaactccagcagccctatacaaca 180  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 158 ggcctatcagctcccaactgcaggctgtgcaagtcacaaactccagcagccctatacaaca 217  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 181 accgacacttcagctggtgcaaggagctgctggtgcaagtcacaaactccagcagcttcg 240  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 218 accgacacttcagctggtgcaaggagctgctggtgcaagtcacaaactccagcagcttcg 277  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 241 tcattggggagaaactgttccacgagtcagctatgagtgagcgtctatctgagaaagc 300  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 278 tcattggggagaaactgttccacgagtcagctatgagtgagcgtctatctgagaaagc 337  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 301 aggtgctgaacttcaccccttgaaagtgctgttccctcactatctgataggttccagcctc 360  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 338 aggtgctgaacttcaccccttgaaagtgctgttccctcactatctgataggttccagcctc 397  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 361 atatcagagagtggtgctcttccctccctggccagctcagcaaacaggctaagcatgtcata 420  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 398 atatcagagagtggtgctcttccctccctggccagctcagcaaacaggctaagcatgtcata 457  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 421 ttgaagtgatgacctgcatatccagagggaattgtgcaaaagctgaaggagcacacagtgaaa 480  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 458 ttgaagtgatgacctgcatatccagagggaattgtgcaaaagctgaaggagcacacagtgaaa 517  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 481 agcttgagagagtgagagatcaaacgaattggaggaactggagctgttcttctatgctc 540  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 518 agcttgagagagtgagagatcaaacgaattggaggaactggagctgttcttctatgctc 577  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 541 tgagaatgctgctgctattgacagagcaaaagctgaaaaatgaataacaccccttcc 600  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 578 tgagaatgctgctgctattgacagagcaaaagctgaaaaatgaataacaccccttcc 637  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 601 ctgctagaataacaattagatgcccccaagcagctattttttaaccacaaaggagaagg 660  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 638 ctgctagaataacaattagatgcccccaagcagctattttttaaccacaaaggagaagg 697  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 661 aagccaaactcccatcatgctgggtggtccaaatgaacccctgcttagttacaaggga 720  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 698 aagccaaactcccatcatgctgggtggtccaaatgaacccctgcttagttacaaggga 757  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 721 aaccaatgcccacttttcttataagaccagaaggtagactttctaaagcatagatattat 780  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 758 aaccaatgcccacttttcttataagaccagaaggtagactttctaaagcatagatattat 817  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 781 tgatacaatttcattgtaactggtgtctctatcacagaaaaaactttatttttaataaa 840  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 818 tgatacaatttcattgtaactggtgtctctatcacagaaaaaactttatttttaataaa 877  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 841 ttgtcttttccataaaaaagatctacttccctcttcccttaggggaaaaaaccccttaata 900  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 878 ttgtcttttccataaaaaagatctacttccctcttcccttaggggaaaaaaccccttaata 937  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 901 gcttcattgtttccataaatcagctactttatattatataatgctattattattattataaga 960  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 938 gcttcattgtttccataaatcagctactttatattatataatgctattattattattataaga 997  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 961 ctgcaattttattatataatcttttataatgctattattattattataagaaacatctcgata 1020  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 998 ctgcaattttattatataatcttttataatgctattattattattataagaaacatctcgata 1057  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 1021 ttgctacttgagtgtaaggctaaatattgataatttatgacaataattatagagctataaca 1080  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 1058 ttgctacttgagtgtaaggctaaatattgataatttatgacaataattatagagctataaca 1117  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 1081 tgtttatttgacctcaataaacaacttggatatcc 1114  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Qy 1118 tgtttatttgacctcaataaacaacttggatatcc 1151  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 6

AAC84310

ID AAC84310 standard; cDNA: 1132 BP.

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

XX AC AAC84310;

PF 19-MAY-2000; 2000WO-US13975.  
 XX 19-MAY-1999; 99US-0134949.  
 PR 15-JUL-1999; 99US-0144270.  
 PR 30-JUL-1999; 99US-0146700.  
 XX 04-OCT-1999; 99US-0157508.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Tang YT, Yue H, Lal P, Burford N, Bandman O, Baughn MR;  
 PI Azlmzal Y, Lu DAM, Patterson C;  
 PI WPI: 2001-025021/03.  
 DR P-PSDB; ABB48074.  
 XX  
 PT New human extracellular signaling nucleic acids and polypeptides useful  
 PT for diagnosing, treating and preventing infections and  
 PT gastrointestinal, neurological, reproductive, and  
 PT autoimmune/inflammatory disorders  
 XX  
 PS Claim 4; Page 108; 114pp; English.  
 XX  
 CC The invention provides human extracellular signaling molecules (EXCS)  
 CC and polynucleotides which identify and encode EXCS. EXCS can be  
 CC expressed by standard recombinant methodology. The amino acid and nucleic  
 CC acid sequences of EXCS are useful for diagnosing, treating and  
 CC preventing infections and gastrointestinal (peptic ulcer, dysphagia,  
 CC pancreatitis), neurological (e.g. epilepsy, ischemic cerebrovascular  
 CC disease, stroke), reproductive (infertility, ovulatory defects,  
 CC endometriosis), autoimmune/inflammatory (actinic keratosis, acquired  
 CC immunodeficiency syndrome (AIDS), Addison's disease), and cell  
 CC proliferative disorders including cancers (of the breast, adrenal gland,  
 CC bone). They may also be used to treat fatal familial insomnia,  
 CC nutritional and metabolic diseases of the nervous system, myopathies,  
 CC mental disorders (anxiety, schizophrenia, mood), as well as infections  
 CC caused by parasites (malaria, leishmania, trypanosoma), viral  
 CC (adenovirus, coronavirus, flavivirus), bacterial (e.g. pneumococcus,  
 CC staphylococcus, bacillus), and fungal (aspergillus, blastomycetes, etc.)  
 CC agonists, pharmaceutical compositions, and antibodies may also be used  
 CC for treating or preventing disorders associated with increased or  
 CC decreased expression or activity of EXCS. EXCS polynucleotides may also  
 CC be used to detect and quantify gene expression in biopsied tissues in  
 CC which expression of EXCS may be correlated with the disease, to determine  
 CC presence or excess expression of EXCS, to monitor regulation of EXCS  
 CC levels during therapeutic intervention, to detect the presence of  
 CC associated disorders, as targets in microarray, to generate hybridization  
 CC probes, and to detect differences in gene sequences among normal, carrier  
 CC or affected individuals. Antibodies may also be used in diagnosing  
 CC disorders, in monitoring patients being treated with EXCS agonists,  
 CC antagonists or inhibitors. Sequences AAC84293-C84318 represent nucleic  
 CC acid molecules encoding the EXCS of the invention.  
 XX  
 SS Sequence 1132 BP; 333 A; 240 C; 226 G; 333 T; 0 other;

Query Match 97.7%; Score 1090; DB 22; Length 1132;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1090; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 tcgagttagaattgttcgaatggccgcctcgcagaaatctgtgagctcttcctattgg 60  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 31 tcgagttagaattgttcgaatggccgcctcgcagaaatctgtgagctcttcctattgg 90  
 QY 61 ggaccttgccaccagctcctcctctcttcttggccctcttgacaggagagcagctg 120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 91 ggaccttgccaccagctcctcctctcttcttggccctcttgacaggagagcagctg 150  
 QY 121 cgcccatcagctcccaactcagcttgacaaagtcacaaactccagcagccctataacca 180  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 151 cgcccatcagctcccaactcagcttgacaaagtcacaaactccagcagccctataacca 210  
 QY 181 accgcaccttcattgctggcgaaggagctagcttggctgatacaacacagcgttcgtc 240  
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 211 accgcaccttcattgctggcgaaggagctagcttggctgatacaacacagcgttcgtc 270  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 241 tcattggggagaaactgttcccgagtgatgagtgagcgtctctctctgataaagc 300  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 271 tcattggggagaaactgttcccgagtgatgagtgagcgtctctctctgataaagc 330  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 301 aggtgctgaacttcaccccttgaagaagtgctgttccctcaatctgatagtttccagcctt 360  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 331 aggtgctgaacttcaccccttgaagaagtgctgttccctcaatctgatagtttccagcctt 390  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 361 atatgcagagagtggtgccccttcctgcagagctcagcaacagggctaaagcacatgcata 420  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 391 atatgcagagagtggtgccccttcctgcagagctcagcaacagggctaaagcacatgcata 450  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 421 ttgaagtgatgacotgcataccagaggaatgtgcaaaagctgaaagacacagtgaaaaa 480  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 451 ttgaagtgatgacotgcataccagaggaatgtgcaaaagctgaaagacacagtgaaaaa 510  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 481 agcttggagagagtgagagatcaaaagcaatggagaaactggatttgcgttcttctctc 540  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 511 agcttggagagagtgagagatcaaaagcaatggagaaactggatttgcgttcttctctc 570  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 541 ttgaagtgatgacotgcataccagaggaatgtgcaaaagctgaaagacacagtgaaaaa 600  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 571 ttgaagtgatgacotgcataccagaggaatgtgcaaaagctgaaagacacagtgaaaaa 630  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 601 ctgctagaaaatacaaatagatgcccccaagcgattttttaaaccacaaagaaagatgg 660  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 631 ctgctagaaaatacaaatagatgcccccaagcgattttttaaaccacaaagaaagatgg 690  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 661 angcacaactccatcatgatgggtgattcccaatgaacccctgcgttagttacaagga 720  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 691 angcacaactccatcatgatgggtgattcccaatgaacccctgcgttagttacaagga 750  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 721 aaccaatgccaactttgtttataagaccagaggtagacttttctaagcatalagatttat 780  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 751 aaccaatgccaactttgtttataagaccagaggtagacttttctaagcatalagatttat 810  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 781 tgatacaatttcattgtaactggtgtctctacacacagaaacaaatatttttaataaa 840  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 811 tgatacaatttcattgtaactggtgtctctacacacagaaacaaatatttttaataaa 870  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 841 ttgtcttttccataaaaaagattacttccattcctttaggggaaaaaacccctaaata 900  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 871 ttgtcttttccataaaaaagattacttccattcctttaggggaaaaaacccctaaata 930  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 901 gtttcattgttccataaatacagctactttatttataaaatgtattttattattataaga 960  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 931 gtttcattgttccataaatacagctactttatttataaaatgtattttattattataaga 990  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 961 ctgcattttattatcatcttttatttaataatggattttattatagaacacatctcgata 1020  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 991 ctgcattttattatcatcttttatttaataatggattttattatagaacacatctcgata 1050  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1021 ttgtctacttgagtgtaaggctaaattgataatttatgacaataattatagagcataaaca 1080  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1051 ttgtctacttgagtgtaaggctaaattgataatttatgacaataattatagagcataaaca 1110  
 QY ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1081 tgtttatttg 1090  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 1111 tgtttatttg 1120  
 DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 7  
 AAF28841  
 ID AAF28841 standard; DNA; 1139 BP.  
 XX  
 AC AAF28841;  
 XX  
 DT 25-APR-2001 (first entry)  
 XX  
 DE Human interleukin-10 homologue IL-D110 DNA sequence.



AAA28839 standard; cDNA; 690 BP.

AAA28839;

04-SEP-2000 (first entry)

Human T cell inducible factor cDNA.

TIF-alpha; T cell derived inducible factor; Interleukin 9; STAT; IL-9; Anti-asthmatic; anti-allergic; cytostatic; inhibitor; antagonist; probe; chromosome 12q15; ss.

Homo sapiens.

Key Location/Qualifiers  
CDS 72..611  
/\*tag= a

W0200024758-A1.

04-MAY-2000.

18-OCT-1999; 99WO-US24424.

26-OCT-1998; 98US-0178973.

16-JUL-1999; 99US-0354243.

(LUDW-) LUDWIG INST CANCER RES.

Dumontier L, Louhed J, Renaud J;

WPI; 2000-422495/36.

P-PSDB; AAY92875.

New nucleic acid molecule encoding a T cell derived inducible factor for treating asthma, an allergy or lymphoma

Claim 1; Page 38; 46pp; English.

This cDNA encodes a human T cell derived inducible factor (TIF). The gene was mapped to chromosome 12q15. The human TIF was identified based on homology to a murine TIF, which was identified by subtraction cloning from a murine lymphoma cell line BW5147 in the presence or absence of interleukin 9 (IL-9). BW5147, can be grown in vitro, without the need to add any cytokines to its culture medium. Many IL-9 activities are mediated by activation of STAT transcription factors. The novel TIFs were expressed in the presence of IL-9, but not in its absence. TIFs induce STAT activation in cells. They can be used, e.g. in the stimulation of regeneration of targeted tissues. Their inhibitors or antagonists can be used to retard, prevent or inhibit differentiation of other tissues. The TIFs and their coding sequences are useful in the treatment of asthma, allergies and lymphoma (claimed). They are also useful for identifying compounds that inhibit or activate T cell induced factor activity in a cell (claimed).

Sequence 690 BP; 182 A; 176 C; 167 G; 165 T; 0 other;

Query Match 57.3%; Score 639; DB 21; Length 690;  
Best Local Similarity 100.0%; Pred. No. 2e-301;  
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 tcgagttagaattgtctgcaatggccgccttcgcaaatctgtgagctcttctctatgg 60  
|||||  
52 tcgagttagaattgtctgcaatggccgccttcgcaaatctgtgagctcttctctatgg 111  
|||||  
61 ggacctggccaccagctgcctctctcttcttggcccttgggtacaggagagcagctg 120  
|||||  
112 ggacctggccaccagctgcctctctcttcttggcccttgggtacaggagagcagctg 171  
|||||  
121 cgccatcagctccctcactgcaggcttgcagagttcgaacttcagcagccttatcacca 180  
|||||  
172 cgccatcagctccctcactgcaggcttgcagagttcgaacttcagcagccttatcacca 231  
|||||

QY 181 accgcaccttcatgctgctaaggaggtgctggtgataacaacacacagcgttcgtc 240  
|||||  
DB 232 accgcaccttcatgctgctaaggaggtgctggtgataacaacacacagcgttcgtc 291  
|||||  
QY 241 tcattggggagaactgttccacggagtcagtatgagtgagcgtgctatctgtatgaagc 300  
|||||  
DB 292 tcattggggagaactgttccacggagtcagtatgagtgagcgtgctatctgtatgaagc 351  
|||||  
QY 301 aggtgctgaacttcacctgaaagaagtgttccctcaatctgaatgaggttccagcctt 360  
|||||  
DB 352 aggtgctgaacttcacctgaaagaagtgttccctcaatctgaatgaggttccagcctt 411  
|||||  
QY 361 atatgcaggaggtgtgtgcttctgcccaggtcagcaacaggcctaagcacatgtcata 420  
|||||  
DB 412 atatgcaggaggtgtgtgcttctgcccaggtcagcaacaggcctaagcacatgtcata 471  
|||||  
QY 421 ttgaagtgatgacctgcatactccagagggaatgtgcataaagctgaaggacacagtgaaaa 480  
|||||  
DB 472 ttgaagtgatgacctgcatactccagagggaatgtgcataaagctgaaggacacagtgaaaa 531  
|||||  
QY 481 agcttgagagagtgagagatcaaacgaattggagaaactgatttctgttattgtctc 540  
|||||  
DB 532 agcttgagagagtgagagatcaaacgaattggagaaactgatttctgttattgtctc 591  
|||||  
QY 541 tgagaaatgcctgctatttgaccagagcaaaagctgaaaaatgaataactaacccccctttcc 600  
|||||  
DB 592 tgagaaatgcctgctatttgaccagagcaaaagctgaaaaatgaataactaacccccctttcc 651  
|||||  
QY 601 ctgctagaataacaattagatgccccaaagcgtatttt 639  
|||||  
DB 652 ctgctagaataacaattagatgccccaaagcgtatttt 690  
|||||

RESULT 9

AAS14875

ID AAS14875 standard; cDNA; 690 BP.

XX

AC AAS14875;

XX

DT 19-DEC-2001 (first entry)

XX

DE Human cDNA encoding T cell derived inducible factor, TIF.

XX

KW Human; T cell derived inducible factor; TIF; ss; antiasthmatic; cytokine; interleukin-9; IL-9; STAT transcription factor; cancer; lymphoma; immune system disorder; allergy; asthma; acquired immunodeficiency syndrome; AIDS; autoimmune diabetes; thyroiditis; melanoma; hepatoma.

KW

XX

OS Homo sapiens.

XX

Key Location/Qualifiers

FT CDS

72..611

FT

/\*tag= a

FT

/product= "TIF"

FT

/transl\_except= (pos:159..161,aa:Glu)

FT

/transl\_except= (pos:378..380,aa:Ile)

FT

/transl\_except= (pos:405..407,aa:Arg)

FT

/transl\_except= (pos:519..521,aa:Cys)

XX

PN US2001024652-A1.

XX

PD 27-SEP-2001.

XX

PF 29-DEC-2000; 2000US-0751797.

XX

PR 18-OCT-1999; 99US-0419568.

XX

PR 26-OCT-1998; 98US-0178973.

XX

PR 16-JUL-1999; 99US-0354243.

XX

PA (DUMO/) DUMOUTIER L.

PA (LOUA/) LOUAHED J.







CC TIFs and their coding sequences are useful in the treatment of asthma,  
CC allergies and lymphoma (claimed). They are also useful for identifying  
CC compounds that inhibit or activate T cell induced factor activity in a  
CC cell (claimed).

XX Sequence 4796 BP; 1339 A; 912 C; 1063 G; 1482 T; 0 other;

Query Match 18.5%; Score 207; DB 21; Length 4796;  
Best Local Similarity 100.0%; Pred. No. 9.6e-91;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tcgaattagaattgtctgcaatggcgccctgcagaaatctgtgagctcttccctatgg 60

Db 52 tcgagttagaattgtctgcaatggcgccctgcagaaatctgtgagctcttccctatgg 111

Qy 61 ggaccctggccaccagctgcctcctcttctgtgcccctctgtgtacaggaggagcagctg 120

Db 112 ggaccctggccaccagctgcctcctcttctgtgcccctctgtgtacaggaggagcagctg 171

Qy 121 cgcctcatcagctcccaactgcaggttgacaagtccaacttccagagccctataccaca 180

Db 172 cgcctcatcagctcccaactgcaggttgacaagtccaacttccagagccctataccaca 231

Qy 181 accgcaccttcagctggcctaaggagg 207

Db 232 accgcaccttcagctggcctaaggagg 258

## RESULT 11

AAAS14876  
ID AAAS14876 standard; DNA; 4797 BP.

XX AC AAAS14876;

DT 19-DEC-2001 (first entry)

DE Human partial genomic DNA for T cell derived inducible factor, TIF.

XX Human; T cell derived inducible factor; TIF; ds; antiallergic;  
KW antiallergic; cytokine; interleukin-9; IL-9; STAT transcription factor;  
KW cancer; lymphoma; immune system disorder; allergy; asthma;  
KW acquired immunodeficiency syndrome; AIDS; autoimmune diabetes;  
KW thyroiditis; melanoma; hepatoma.

XX Homo sapiens.

XX US2001024652-A1.

XX 27-SEP-2001.

XX 29-DEC-2000; 2000US-0751797.

XX 18-OCT-1999; 99US-0419568.

XX 26-OCT-1998; 98US-0178973.

XX 16-JUL-1999; 99US-0354243.

XX (DUMO/) DUMOUTIER L.

XX (LOUA/) LOUAHED J.

XX (RENA/) RENAULD J.

XX Dumoutier L, Louahed J, Renauld J;

XX WPI; 2001-638496/73.

XX New isolated nucleic acid molecules encoding T cell inducible factors,  
PT useful as markers for expression or effect of interleukin (IL)-9 in a  
PT subject and diagnosing susceptibility to asthma or allergy

XX Claim 1; Page 17-19; 26pp; English.

XX The invention relates to an isolated nucleic acid molecule, which encodes  
CC a T cell derived inducible factor (TIF) which are upregulated by the

CC cytokine interleukin-9 (IL-9) and induce STAT transcription factor  
CC activation. The TIF proteins (or their mutants) may be used to test IL-9  
CC antagonists for their potency against lymphomas, immune system  
CC disorders, allergies, asthma, acquired immunodeficiency syndrome (AIDS),  
CC autoimmune diabetes and thyroiditis. TIF molecules promote regeneration  
CC or inhibit differentiation of tissue types in which they are active and  
CC therefore be used to develop treatments for melanomas and hepatomas.  
CC The present sequence a partial genomic sequence for Human TIF.

XX Sequence 4797 BP; 1339 A; 910 C; 1063 G; 1485 T; 0 other;

Query Match 18.5%; Score 207; DB 22; Length 4797;  
Best Local Similarity 100.0%; Pred. No. 9.6e-91;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tcgaattagaattgtctgcaatggcgccctgcagaaatctgtgagctcttccctatgg 60

Db 52 tcgagttagaattgtctgcaatggcgccctgcagaaatctgtgagctcttccctatgg 111

Qy 61 ggaccctggccaccagctgcctcctcttctgtgcccctctgtgtacaggaggagcagctg 120

Db 112 ggaccctggccaccagctgcctcctcttctgtgcccctctgtgtacaggaggagcagctg 171

Qy 121 cgcctcatcagctcccaactgcaggttgacaagtccaacttccagagccctataccaca 180

Db 172 cgcctcatcagctcccaactgcaggttgacaagtccaacttccagagccctataccaca 231

Qy 181 accgcaccttcagctggcctaaggagg 207

Db 232 accgcaccttcagctggcctaaggagg 258

## RESULT 12

AAAD09746

ID AAAD09746 standard; cDNA; 778 BP.

XX AC AAAD09746;

DT 10-SEP-2001 (first entry)

DE Mouse ZCYTO18 cDNA.

XX Mouse; cytostatic; cytokine; ZCYTO18 protein; genetic abnormality;  
KW cancer; inflammation; gene therapy; ss.

XX Mus musculus.

XX Key Location/Qualifiers

FT CDS 47..586

FT /\*tag= a

FT /product= "Mouse ZCYTO18 protein"

FT sig\_peptide 47..145

FT /\*tag= b

FT mat\_peptide 146..583

FT /\*tag= c

FT /product= "Mouse mature ZCYTO18 protein"

XX WO200146422-A1.

XX 28-JUN-2001.

XX 22-DEC-2000; 2000WO-US35308.

XX 23-DEC-1999; 99US-0471767.

XX 01-DEC-2000; 2000US-0250841.

XX (ZYMO ) ZYMOGENETICS INC.

XX Presnell SR, Kindsvogel W;

XX WPI; 2001-408648/43.

DR P-PSDB; AAE05052.

XX Novel human cytokine polypeptide, ZCYTO18, useful for treating cancer -  
 PT Example 5; Page 160-162; 167pp; English.  
 PS  
 XX The patent discloses novel human cytokine, ZCYTO18 protein and its  
 CC corresponding DNA. ZCYTO18 protein induces proliferation of cells  
 CC expressing zcytor11, a receptor for ZCYTO18 or induces cytotoxicity.  
 CC In K562 cells, ZCYTO18 DNA is useful for detecting a genetic  
 CC abnormality in a patient. ZCYTO18 DNA and its antibodies are useful  
 CC for detecting cancer and inflammation. ZCYTO18 protein is useful for  
 CC killing cancer cells. It is useful for increasing platelets in a  
 CC patient or injured tissue. It is also used in gene therapy.  
 CC The present sequence is a cDNA encoding mouse cytokine, ZCYTO18.  
 XX  
 SQ Sequence 778 BP; 213 A; 193 C; 197 G; 175 T; 0 other;

Query Match 3.1%; Score 35; DB 22; Length 778;  
 Best Local Similarity 100.0%; Pred. No. 7e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 471 acagtgaagcttgagagagtgaggagatcaa 505  
 |||||  
 Db 497 acagtgaagcttgagagagtgaggagatcaa 531  
 |||||

RESULT 13  
 AAA28815  
 ID AAA28815 standard; cDNA; 1119 BP.  
 AC AAA28815;  
 DT 04-SEP-2000 (first entry)  
 XX Murine T cell inducible factor alpha cDNA.  
 XX  
 XX TIF-alpha; T cell derived inducible factor; Interleukin 9; STAT; IL-9;  
 KW Anti-asthmatic; anti-allergic; cytostatic; inhibitor; antagonist; ss.  
 KW  
 OS Mus sp.  
 XX  
 XX Key Location/Qualifiers  
 FH 52..591  
 FT /\*tag= a  
 FT primer\_bind  
 FT /note= "PCR primer appearing as AAS14861"  
 FT complement (764..784)  
 FT /\*tag= c  
 FT /note= "PCR primer appearing as AAS14862"  
 PN WO200024758-A1.  
 XX  
 XX 04-MAY-2000.  
 PD  
 XX 18-OCT-1999; 99WO-US24424.  
 XX  
 XX 26-OCT-1998; 98US-0178973.  
 PR 16-JUL-1999; 99US-0354243.  
 XX  
 XX {LUDW-} LUDWIG INST CANCER RES.  
 PA  
 XX Dumoutier L, Louhed J, Renaud J;  
 PI  
 XX WPI; 2000-422495/36.  
 DR P-PSDB; AAY92877.  
 XX  
 XX New nucleic acid molecule encoding a T cell derived inducible factor  
 PT for treating asthma, an allergy or lymphoma  
 PT  
 XX Claim 1; Page 33; 46pp; English.  
 PS  
 XX This cDNA encodes T cell derived inducible factor (TIF) alpha identified  
 CC by subtraction cloning from a murine lymphoma cell line BW5147 in the  
 CC presence or absence of interleukin 9 (IL-9). BW5147, can be grown in  
 CC vitro, without the need to add any cytokines to its culture medium. Many  
 CC IL-9 activities are mediated by activation of STAT transcription  
 CC factors. The novel TIFs were expressed in the presence of IL-9, but not

CC in its absence. TIFs induce STAT activation in cells. They can be used,  
 CC e.g. in the stimulation of regeneration of targeted tissues. Their  
 CC inhibitors or antagonists can be used to retard, prevent or inhibit  
 CC differentiation of other tissues. The TIFs and their coding sequences are  
 CC useful in the treatment of asthma, allergies and lymphoma (claimed). They  
 CC are also useful for identifying compounds that inhibit or activate T cell  
 CC induced factor activity in a cell (claimed).  
 XX  
 SQ Sequence 1119 BP; 352 A; 233 C; 232 G; 302 T; 0 other;

Query Match 3.1%; Score 35; DB 21; Length 1119;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 471 acagtgaagcttgagagagtgaggagatcaa 505  
 |||||  
 Db 502 acagtgaagcttgagagagtgaggagatcaa 536  
 |||||

RESULT 14  
 AAS14858  
 ID AAS14858 standard; cDNA; 1119 BP.  
 AC AAS14858;  
 XX  
 XX 19-DEC-2001 (first entry)  
 XX  
 XX Mouse cDNA encoding T cell derived inducible factor, TIFalpha.  
 DE  
 XX  
 XX Mouse; T cell derived inducible factor; TIFalpha; ss; anti-allergic;  
 KW antiasthmatic; cytokine; interleukin-9; IL-9; STAT transcription factor;  
 KW cancer; lymphoma; immune system disorder; allergy; asthma;  
 KW acquired immunodeficiency syndrome; AIDS; autoimmune diabetes;  
 KW thyroiditis; melanoma; hepatoma.  
 XX  
 XX Mus musculus.  
 XX  
 XX Key Location/Qualifiers  
 FH 52..591  
 FT /\*tag= a  
 FT /product= "TIFalpha"  
 FT 106..126  
 FT /\*tag= b  
 FT primer\_bind  
 FT /note= "PCR primer appearing as AAS14861"  
 FT complement (764..784)  
 FT /\*tag= c  
 FT /note= "PCR primer appearing as AAS14862"  
 PN US2001024652-A1.  
 XX  
 XX 27-SEP-2001.  
 PD  
 XX 29-DEC-2000; 2000US-0751797.  
 XX  
 XX 18-OCT-1999; 99US-0419568.  
 PR 26-OCT-1998; 98US-0178973.  
 PR 16-JUL-1999; 99US-0354243.  
 XX  
 XX {DUMO/} DUMOUTIER L.  
 PA {LOUA/} LOUAHED J.  
 PA {RENA/} RENAULD J.  
 XX  
 XX Dumoutier L, Louhed J, Renaud J;  
 PI  
 XX WPI; 2001-638496/73.  
 DR P-PSDB; AAU09090.  
 XX  
 XX New isolated nucleic acid molecules encoding T cell inducible factors,  
 PT useful as markers for expression or effect of interleukin (IL)-9 in a  
 PT subject and diagnosing susceptibility to asthma or allergy -  
 XX Claim 1; Page 10; 26pp; English.  
 PS

XX The invention relates to an isolated nucleic acid molecule, which encodes  
 CC a T cell derived inducible factor (TIF) which are upregulated by the  
 CC cytokine interleukin-9 (IL-9) and induce STAT transcription factor  
 CC activation. The TIF proteins (or their mutants) may be used to test IL-9  
 CC ant/agonists for their potency against lymphomas, immune system  
 CC disorders, allergies, asthma, acquired immunodeficiency syndrome (AIDS),  
 CC autoimmune diabetes and thyroiditis. TIF molecules promote regeneration  
 CC or inhibit differentiation of tissue types in which they are active and  
 CC therefore be used to develop treatments for melanomas and hepatomas.  
 CC The present sequence encodes mouse Tifalpha.  
 XX  
 XX Sequence 1119 BP; 352 A; 233 C; 232 G; 302 T; 0 other;

Query Match 3.1%; Score 35; DB 22; Length 1119;  
 Best Local Similarity 100.0%; Pred. No. 7.1e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 acagtgaagaaagcttgagagagtgagagatcaa 505  
 |||||  
 Db 502 acagtgaagaaagcttgagagagtgagagatcaa 536

## RESULT 15

AAC81774  
 ID AAC81774 standard; cDNA; 1166 BP.

XX  
 AC AAC81774;

XX  
 DT 23-FEB-2001 (first entry)

XX  
 DE Murine GIL-19 protein coding sequence.

XX  
 KW Mouse; GIL-19/AE289; IL-10; interleukin-10; nutrition;

XX  
 KW cell proliferation; immune stimulation; immune suppression;

XX  
 KW haematopoiesis regulation; tissue growth; inflammation; cancer; ss.

XX  
 OS Mus sp.

XX  
 PN WO200065027-A2.

XX  
 PD 02-NOV-2000.

XX  
 PF 28-APR-2000; 2000WO-US11479.

XX  
 PR 28-APR-1999; 99US-0131473.

XX  
 PA (GENY ) GENETICS INST INC.

XX  
 PI Jacobs K, Fouser L, Spaulding V, Xuan D;

XX  
 DR WPI; 2000-687325/67.

XX  
 PT Human GIL-19 protein that shows a high degree of homology to IL  
 PT (interleukin)-10, useful in upregulation of humoral immune responses,  
 PT as an antiinflammatory agent and as a modulator of immune responses  
 PT associated with injury .

XX  
 PS Disclosure; Fig 1; 60pp; English.

XX  
 CC The present invention provides the protein and coding sequences for the  
 CC novel human GIL-19/AE289 protein. The protein shows homology to  
 CC interleukin-10 (IL-10) and is assumed to be a cytokine. It can be used  
 CC in the regulation of cell proliferation and differentiation,  
 CC haematopoiesis, immune stimulation or suppression, tissue growth and  
 CC tumour inhibition. In addition, it also has uses in the treatment of  
 CC inflammation and in nutrition.

XX  
 SQ Sequence 1166 BP; 375 A; 241 C; 244 G; 306 T; 0 other;

Query Match 3.1%; Score 35; DB 21; Length 1166;

Best Local Similarity 100.0%; Pred. No. 7.1e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 acagtgaagaaagcttgagagagtgagagatcaa 505  
 |||||  
 Db 525 acagtgaagaaagcttgagagagtgagagatcaa 559

Search completed: September 23, 2002, 02:11:28  
 Job time: 5249 sec

us-09-746-375-1.rng

Mon Sep 23 09:43:52 2002

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2002, 00:51:20 ; Search time 51.77 Seconds  
(without alignments)  
5295.093 Million cell updates/sec

Title: us-09-746-375-1

Perfect score: 1116

Sequence: 1 tcgagtagaattgtctgca.....ataaacacttgatccta 1116

Scoring table:

OLIGO\_NUC

Gapop 50.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 14

Total number of hits satisfying chosen parameters: 1870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued\_Patents\_NA.\*

- 1: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/1na/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/1na/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/1na/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/1na/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/1na/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	639	57.3	690	4	US-09-354-243B-24
2	207	18.5	4797	4	US-09-354-243B-25
3	35	3.1	1119	4	US-09-178-973B-7
4	35	3.1	1119	4	US-09-354-243B-7
5	31	2.8	7445	4	US-09-178-973B-8
6	31	2.8	7445	4	US-09-354-243B-8
7	28	2.5	1111	4	US-09-178-973B-9
8	28	2.5	1111	4	US-09-354-243B-9
9	28	2.5	5935	4	US-09-178-973B-17
10	28	2.5	5935	4	US-09-354-243B-29
11	22	2.0	22	4	US-09-354-243B-16
12	21	1.9	21	4	US-09-354-243B-18
13	21	1.9	21	4	US-09-354-243B-19
14	20	1.8	2908	3	US-08-487-799-1
15	20	1.8	22846	2	US-08-469-461-3
16	20	1.8	22846	3	US-07-890-609-3
17	18	1.6	1939	1	US-07-715-751B-2
18	17	1.5	356	4	US-09-319-730-16
19	17	1.5	452	2	US-08-378-235B-4
20	17	1.5	573	2	US-08-290-665A-134
21	17	1.5	573	5	PCT-US95-10398-134
22	17	1.5	1519	1	US-07-971-759-19
23	17	1.5	1628	2	US-08-883-515-3
24	17	1.5	3217	4	US-09-232-200-64
25	17	1.5	3217	4	US-09-232-197-64
26	17	1.5	3217	4	US-09-232-201-64
27	17	1.5	9620	4	US-08-952-127-11

#### ALIGNMENTS

##### RESULT 1

US-09-354-243B-24  
; Sequence 24, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louheid, Jamila  
; APPLICANT: Renaud, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
; TITLE OF INVENTION: (TIFs)  
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354,243B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: US09/178,973  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 24  
; LENGTH: 690  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-354-243B-24

Query Match 57.3%; Score 639; DB 4; Length 690;  
Best Local Similarity 100.0%; Pred. No. 5e-302;  
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	1	tcgagttagaattgtctgcaatggccgccttcgagaaatctgtgagctcttcttattgg	60
Db	52	tcgagttagaattgtctgcaatggccgccttcgagaaatctgtgagctcttcttattgg	111
Oy	61	ggacctgggacacagctgcctctcttcttggcctcttggtagaggagagcagctg	120
Db	112	ggacctgggacacacagctgcctctcttcttggcctcttggtagaggagagcagctg	171
Oy	121	cgcccatcagctcccaactgcaggcttgacaagctccaaactccagcagcctatatacca	180
Db	172	cgcccatcagctcccaactgcaggcttgacaagctccaaactccagcagcctatatacca	231
Oy	181	accgacacttaactgtgctgaaggaggttagcttggctgataaacacacagcgttcg	240
Db	232	accgacacttaactgtgctgaaggaggttagcttggctgataaacacacagcgttcg	291
Oy	241	tcattggggagaaaactgttcacacagtcagtcagtcagtcagtcagtcagtcagtc	300
Db	292	tcattggggagaaaactgttcacacagtcagtcagtcagtcagtcagtcagtcagtc	351
Oy	301	agggtgtaacttcaccttggaagtgctgttccctcaatctgtaggtttccagcctt	360

Db 352 aggtgtaacttcaccccttgagaagtgcttccctcaatctgataggtccagcctt 411  
Qy 361 atatcagagagggtggtcccttctgcccaggtccagcaacaggtcaagcacatgtcata 420  
Db 412 atatcagagagggtggtcccttctgcccaggtccagcaacaggtcaagcacatgtcata 471  
Qy 421 ttgaagtgatgacctgcatactacagagggaatgtgcaaaagctggaaggacacagtgaaaa 480  
Db 472 ttgaagtgatgacctgcatactacagagggaatgtgcaaaagctggaaggacacagtgaaaa 531  
Qy 481 agcttgagagagtgagagagatcaaaagcaattggagaactggattgtctttatgtctc 540  
Db 532 agcttgagagagtgagagagatcaaaagcaattggagaactggattgtctttatgtctc 591  
Qy 541 tgagaaatgctgcatttgccagagcaaaagctgaaaaatgaataaactaacccctttcc 600  
Db 592 tgagaaatgctgcatttgccagagcaaaagctgaaaaatgaataaactaacccctttcc 651  
Qy 601 ctgctagaataaacaattagatgcccaagcgattttt 639  
Db 652 ctgctagaataaacaattagatgcccaagcgattttt 690

RESULT 2  
US-09-354-243B-25  
; Sequence 25, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fa  
; TITLE OF INVENTION: (Tifs)  
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354,243B  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 25  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-354-243B-25

Query Match 18.5%; Score 207; DB 4; Length 4797;  
Best Local Similarity 100.0%; Pred. No. 1.6e-91;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 tcagattagaattgtctgcaatggccgccttcgagaaatctgtgagctcttcttcattgg 60  
Db 52 tcagattagaattgtctgcaatggccgccttcgagaaatctgtgagctcttcttcattgg 111  
Qy 61 ggaacctggccacagctgctctctctcttggccctctgttacaggagagagcagctg 120  
Db 112 ggaacctggccacagctgctctctctcttggccctctgttacaggagagagcagctg 171  
Qy 121 gcccattcagctcccaactgcaggcttgacaagtcccaacttccagcagccctatatcaca 180  
Db 172 gcccattcagctcccaactgcaggcttgacaagtcccaacttccagcagccctatatcaca 231  
Qy 181 accgacattcagctgctgctaaggagg 207  
Db 232 accgacattcagctgctgctaaggagg 258

RESULT 3  
US-09-178-973B-7  
; Sequence 7, Application US/09178973B

; Patent No. 6274710  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renaud, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
; TITLE OF INVENTION: (Tifs)  
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
; FILE REFERENCE: LUD 5543  
; CURRENT APPLICATION NUMBER: US/09/178,973B  
; CURRENT FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 7  
; LENGTH: 1119  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-178-973B-7

Query Match 3.1%; Score 35; DB 4; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 acagtgaagaagcttgagagagtgaggagagatcaa 505  
Db 502 acagtgaagaagcttgagagagtgaggagagatcaa 536

RESULT 4  
US-09-354-243B-7  
; Sequence 7, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renaud, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
; TITLE OF INVENTION: (Tifs)  
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354,243B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 7  
; LENGTH: 1119  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
US-09-354-243B-7

Query Match 3.1%; Score 35; DB 4; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 acagtgaagaagcttgagagagtgaggagagatcaa 505  
Db 502 acagtgaagaagcttgagagagtgaggagagatcaa 536

RESULT 5  
US-09-178-973B-8  
; Sequence 8, Application US/09178973B  
; Patent No. 6274710  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Renaud, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
; TITLE OF INVENTION: (Tifs)  
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof

FILE REFERENCE: LUD 5543  
CURRENT APPLICATION NUMBER: US/09/178,973B  
CURRENT FILING DATE: 1998-10-26  
NUMBER OF SEQ ID NOS: 17  
SEQ ID NO 8  
LENGTH: 7445  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-178-973B-8

Query Match 2.8%; Score 31; DB 4; Length 7445;  
Best Local Similarity 100.0%; Pred. No. 8.9e-06;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 ttatttaataatggtattttattatagaacat 1011  
|||||  
DB 7008 ttatttaataatggtattttattatagaacat 7038

RESULT 6  
US-09-354-243B-8  
Sequence 8, Application US/09354243B  
Patent No. 6359117  
GENERAL INFORMATION:  
APPLICANT: Dumoutier, Laure  
APPLICANT: Louhed, Jamila  
APPLICANT: Renaud, Jean-Christophe  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac  
TITLE OF INVENTION: (TIFS)  
TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
FILE REFERENCE: LUD 5543.1  
CURRENT APPLICATION NUMBER: US/09/354,243B  
CURRENT FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: US09/178,973  
PRIOR FILING DATE: 1998-10-26  
NUMBER OF SEQ ID NOS: 29  
SEQ ID NO 8  
LENGTH: 7445  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
US-09-354-243B-8

Query Match 2.8%; Score 31; DB 4; Length 7445;  
Best Local Similarity 100.0%; Pred. No. 8.9e-06;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 ttatttaataatggtattttattatagaacat 1011  
|||||  
DB 7008 ttatttaataatggtattttattatagaacat 7038

RESULT 7  
US-09-178-973B-9  
Sequence 9, Application US/09178973B  
Patent No. 6274710  
GENERAL INFORMATION:  
APPLICANT: Dumoutier, Laure  
APPLICANT: Louhed, Jamila  
APPLICANT: Renaud, Jean-Christophe  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac  
TITLE OF INVENTION: (TIFS)  
TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
FILE REFERENCE: LUD 5543  
CURRENT APPLICATION NUMBER: US/09/178,973B  
CURRENT FILING DATE: 1998-10-26  
NUMBER OF SEQ ID NOS: 17  
SEQ ID NO 9  
LENGTH: 1111  
TYPE: DNA  
ORGANISM: Mus musculus

US-09-178-973B-9

Query Match 2.5%; Score 28; DB 4; Length 1111;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 ttatttaataatggtattttattatagaacat 1008  
|||||  
DB 981 ttatttaataatggtattttattatagaacat 1008

RESULT 8  
US-09-354-243B-9  
Sequence 9, Application US/09354243B  
Patent No. 6359117  
GENERAL INFORMATION:  
APPLICANT: Dumoutier, Laure  
APPLICANT: Louhed, Jamila  
APPLICANT: Renaud, Jean-Christophe  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
TITLE OF INVENTION: (TIFS)  
TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
FILE REFERENCE: LUD 5543.1  
CURRENT APPLICATION NUMBER: US/09/354,243B  
CURRENT FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: US09/178,973  
PRIOR FILING DATE: 1998-10-26  
NUMBER OF SEQ ID NOS: 29  
SEQ ID NO 9  
LENGTH: 1111  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
US-09-354-243B-9

Query Match 2.5%; Score 28; DB 4; Length 1111;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 ttatttaataatggtattttattatagaacat 1008  
|||||  
DB 981 ttatttaataatggtattttattatagaacat 1008

RESULT 9  
US-09-178-973B-17  
Sequence 17, Application US/09178973B  
Patent No. 6274710  
GENERAL INFORMATION:  
APPLICANT: Dumoutier, Laure  
APPLICANT: Louhed, Jamila  
APPLICANT: Renaud, Jean-Christophe  
TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
TITLE OF INVENTION: (TIFS)  
TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
FILE REFERENCE: LUD 5543  
CURRENT APPLICATION NUMBER: US/09/178,973B  
CURRENT FILING DATE: 1998-10-26  
NUMBER OF SEQ ID NOS: 17  
SEQ ID NO 17  
LENGTH: 5935  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-178-973B-17

Query Match 2.5%; Score 28; DB 4; Length 5935;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 ttatttaataatggtattttattatagaacat 1008

Db 5694 tttattaataggattattattagaaa 5721  
|||||

RESULT 10  
US-09-354-243B-29  
; Sequence 29, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renauld, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Peptides (TIFs)  
; TITLE OF INVENTION: (TIFs)  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354,243B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: US09/178,973  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 29  
; LENGTH: 5935  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-354-243B-29

Query Match 2.5%; Score 28; DB 4; Length 5935;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 981 tttattaataggattattattagaaa 1008  
|||||

Db 5694 tttattaataggattattattagaaa 5721  
|||||

RESULT 11  
US-09-354-243B-16/c  
; Sequence 16, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renauld, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Peptides (TIFs)  
; TITLE OF INVENTION: (TIFs)  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354,243B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: US09/178,973  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 16  
; LENGTH: 22  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-354-243B-16

Query Match 2.0%; Score 22; DB 4; Length 22;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 tgaataagcttgagagagtg 496  
|||||

Db 22 TGAATAAGCTTGAGAGAGTGG 1

US-09-354-243B-18/c  
; Sequence 18, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renauld, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Peptides (TIFs)  
; TITLE OF INVENTION: (TIFs)  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354,243B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: US09/178,973  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 18  
; LENGTH: 21  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-354-243B-18

Query Match 1.9%; Score 21; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 370 agtggtgccccttcttgccca 390  
|||||

Db 21 AGTGTGTCCTTCTTGCCCA 1  
|||||

RESULT 13  
US-09-354-243B-19/c  
; Sequence 19, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renauld, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Peptides (TIFs)  
; TITLE OF INVENTION: (TIFs)  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354,243B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: US09/178,973  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 19  
; LENGTH: 21  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-354-243B-19

Query Match 1.9%; Score 21; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.67;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 370 agtggtgccccttcttgccca 390  
|||||

Db 21 AGTGTGTCCTTCTTGCCCA 1  
|||||

RESULT 14  
US-08-487-799-1  
; Sequence 1, Application US/08487799C  
; Patent No. 6010908  
; GENERAL INFORMATION:  
; APPLICANT: Gruenert, Dieter C.



```

; APPLICANT: Kunzelmann, Karl
; TITLE OF INVENTION: GENE THERAPY BY SMALL FRAGMENTS HOMOLOGOUS REPLACEMENT
; FILE REFERENCE: 480.18-1(HV)
; CURRENT APPLICATION NUMBER: US/08/487,799C
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/933,471
; EARLIER FILING DATE: 1992-08-21
; EARLIER APPLICATION NUMBER: 08/409,544
; EARLIER FILING DATE: 1995-03-24
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: human
US-08-487-799-1

```

```

Query Match      1.8%; Score 20; DB 3; Length 2908;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 822 caattattttttaataat 841
Db 1453 caattattttttaataat 1472

```

```

RESULT 15
US-08-469-461-3
; Sequence 3, Application US/08469461B
; Patent No. 5981178
; GENERAL INFORMATION:
; APPLICANT: Tsui, Lap-Chee
; APPLICANT: Rommins, Johanna M.
; APPLICANT: Kerem, Bat-Sheva
; TITLE OF INVENTION: Introns and Exons of the Cystic Fibrosis Gene and
; FILE REFERENCE: 3477-61, 033477/139840
; CURRENT APPLICATION NUMBER: US/08/469,461B
; CURRENT FILING DATE: 1995-06-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 22846
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-469-461-3

```

```

Query Match      1.8%; Score 20; DB 2; Length 22846;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 822 caattattttttaataat 841
Db 9242 caattattttttaataat 9261

```

Search completed: September 23, 2002, 02:49:51  
Job time: 7111 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2002, 00:42:04 : Search time 1730.87 Seconds  
(without alignments)  
8702.330 Million cell updates/sec

Title: US-09-746-375-1  
Perfect score: 1116  
Sequence: 1 tcgagtagaattgtctgca.....ataaacacttgatataccta 1116

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 14

Total number of hits satisfying chosen parameters: 158874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_gss:\*
- 13: em\_gss\_hum:\*
- 14: em\_gss\_inv:\*
- 15: em\_gss\_pln:\*
- 16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	66	5.9	389	12	AQ104025
C 2	28	2.5	256	9	AV030414
C 3	28	2.5	700	11	AK005228
C 4	23	2.1	116	10	BI314237
C 5	22	2.0	381	10	BI341427
C 6	22	2.0	431	10	BI341431
C 7	22	2.0	446	12	A2167394
C 8	22	2.0	718	12	A0780852
C 9	21	1.9	182	10	W21516
C 10	21	1.9	192	10	W21016
C 11	21	1.9	210	9	AA341708
C 12	21	1.9	234	9	AA334932
C 13	21	1.9	240	9	AW486359
C 14	21	1.9	251	10	W05048
C 15	21	1.9	262	9	AA331022
C 16	21	1.9	295	9	AV097026
C 17	21	1.9	301	10	BF827895

C 18	21	1.9	304	10	F00991
C 19	21	1.9	305	9	AA486360
C 20	21	1.9	315	9	AA346165
C 21	21	1.9	327	10	BF827890
C 22	21	1.9	335	9	AA328929
C 23	21	1.9	378	9	AA112190
C 24	21	1.9	384	12	A2223713
C 25	21	1.9	391	9	AA180126
C 26	21	1.9	401	9	AA194504
C 27	21	1.9	401	9	AA346264
C 28	21	1.9	422	10	BM276287
C 29	21	1.9	429	10	W88590
C 30	21	1.9	433	9	AA192403
C 31	21	1.9	458	9	AA182400
C 32	21	1.9	464	9	AA196108
C 33	21	1.9	473	10	BM276043
C 34	21	1.9	476	12	AQ203627
C 35	21	1.9	478	12	AQ664172
C 36	21	1.9	479	9	AW104304
C 37	21	1.9	480	10	W93616
C 38	21	1.9	503	12	AZ478642
C 39	21	1.9	507	12	AQ992968
C 40	21	1.9	529	10	BI816005
C 41	21	1.9	545	10	BM276174
C 42	21	1.9	548	10	BM275812
C 43	21	1.9	549	10	BM275687
C 44	21	1.9	551	10	BM275051
C 45	21	1.9	555	10	BM274386

ALIGNMENTS

RESULT 1  
AQ104025/c  
LOCUS  
DEFINITION  
HS\_3108\_B1\_C01\_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3108 Col=1 Row=F, DNA sequence.  
ACCESSION  
AQ104025  
VERSION  
AQ104025.1  
KEYWORDS  
GI:3478961  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 389)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,K., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
TITLE  
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
JOURNAL  
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
MEDLINE  
99380589  
COMMENT  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 3108 row: F column: 1  
Class: BAC ends  
High quality sequence stop: 389.  
Location/Qualifiers  
1. 389  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="plate=3108 Col=1 Row=F"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 119 a 86 c 65 g 118 t 1 others  
ORIGIN

Query Match 5.9%; Score 66; DB 12; Length 389;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 417 catatgaaggtgatccctgcattatccagaggaatgtgcaaaagctgaaggacacagt 476  
|||||  
Db 144 CATATTGAAGTGTGATGACCTGCATATCCAGAGGAGTGTGCAAAAGCTGAAGGACACAGT 85  
QY 477 aaaaag 482  
|||||  
Db 84 AAAAAG 79

RESULT 2  
AV030414 AV030414 256 bp mRNA linear EST 31-AUG-1999  
LOCUS  
DEFINITION AV030414 Mus musculus adult C57BL/6J cerebellum Mus musculus cDNA  
clone 1500012D04, mRNA sequence.  
ACCESSION AV030414  
VERSION AV030414.1 GI:4829962  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 256)  
AUTHORS Carninci, P., Shibata, K., Ozawa, Y., Konno, H., Itoh, M., Aizawa, K.,  
Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara,  
A., Hayatsu, N., Kojima, Y., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J.,  
Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C.,  
Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara,  
Y., Suzuki, H., Suzuki, H., Tatenos, M., Tomaru, Y., Tomimaga, N.,  
Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M.,  
Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.  
RIKEN Mouse ESTs  
Unpublished (1999)  
Contact: Chie Owa  
Genome Science Laboratory  
RIKEN  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-9145  
Fax: 81-298-36-9098  
Email: genome-res@rtc.riken.go.jp  
Thermolabile and thermostable enzymes by  
trehalose and its application for the synthesis of full length cDNA  
(Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
Transcriptional sequencing: A method for DNA sequencing using RNA  
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

FEATURES  
Source  
1. 256  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="1500012D04"  
/clone\_lib="Mus musculus adult C57BL/6J cerebellum"  
/sex="male"  
/tissue\_type="cerebellum"  
/dev\_stage="adult"  
88 a 29 c 41 g 98 t

Query Match 2.5%; Score 28; DB 9; Length 256;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 981 ttattatgaatgattattattatagaaa 1008

Db 131 TTTATTAAATATGCGATTATTATTATAGAA 158  
|||||

RESULT 3  
AK005228 AK005228 700 bp mRNA linear HTC 19-JAN-2002  
LOCUS  
DEFINITION Mus musculus adult male cerebellum cDNA, RIKEN full-length enriched  
library, clone:1500012D04:Interleukin 10-related T cell-derived  
inducible factor, full insert sequence.  
ACCESSION AK005228  
VERSION AK005228.1 GI:12837639  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (strain:C57BL/6J) adult male cerebellum cDNA to mRNA,  
clone\_lib:RIKEN full-length enriched mouse cDNA library  
clone:1500012D04.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (sites)  
AUTHORS Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2 (sites)  
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159  
REFERENCE 3 (sites)  
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
TITLE RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multichannel sequencer  
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
MEDLINE 20530913  
PUBMED 11076861  
REFERENCE 4 (sites)  
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
TITLE Functional annotation of a full-length mouse cDNA collection  
JOURNAL Nature 409, 685-690 (2001)  
REFERENCE 5 (bases 1 to 700)  
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,  
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,  
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,  
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Queckenbush, J.,  
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yamanaka, I.,  
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@rtc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222.

Fax:91-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGAGAGAGTCCAGAGCTCTTTTCTTTTNN 3', cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 5' GAGAGAGAGCGCCGCAATATTCGAGTAATTAATTAATCCCCCCC 3', cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI. Host: SOLR.

Location/Qualifiers  
1. 700  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="MGD:MGI:1900981"  
/db\_xref="taxon:10090"  
/clone="1500012D04"  
/sex="male"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
1. 700  
/gene="Iltif"  
1. 700  
/gene="Iltif"  
/note="data source:MGD, source key:MGI:1355307, evidence:ISS  
Interleukin 10-related T cell-derived inducible factor"  
255 a 105 c 133 g 207 t

Query Match 2.5%; Score 28; DB 11; Length 700;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 ttatttaattggtatttattataga 1008  
|||||  
Db 572 TTTTATTATGATGATTTATTATAGAA 599  
|||||

RESULT 4  
BI314237  
LOCUS dai33c12.xl NICHG XGC Sp1 Xenopus laevis cDNA clone IMAGE:4963582  
DEFINITION 3', mRNA sequence.  
ACCESSION BI314237  
VERSION BI314237.1 GI:14988564  
KEYWORDS EST.  
SOURCE African clawed frog.  
ORGANISM Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus.  
1 (bases 1 to 116)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps@email.nih.gov  
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LNL at: [info@image.llnl.gov](mailto:info@image.llnl.gov)  
Seq primer: -400P from Gibco.

FEATURES  
source  
Location/Qualifiers  
1. 700  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="MGD:MGI:1900981"  
/db\_xref="taxon:10090"  
/clone="1500012D04"  
/sex="male"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
1. 700  
/gene="Iltif"  
1. 700  
/gene="Iltif"  
/note="data source:MGD, source key:MGI:1355307, evidence:ISS  
Interleukin 10-related T cell-derived inducible factor"  
255 a 105 c 133 g 207 t

Query Match 2.5%; Score 28; DB 11; Length 700;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 ttatttaattggtatttattataga 1008  
|||||  
Db 572 TTTTATTATGATGATTTATTATAGAA 599  
|||||

RESULT 4  
BI314237  
LOCUS dai33c12.xl NICHG XGC Sp1 Xenopus laevis cDNA clone IMAGE:4963582  
DEFINITION 3', mRNA sequence.  
ACCESSION BI314237  
VERSION BI314237.1 GI:14988564  
KEYWORDS EST.  
SOURCE African clawed frog.  
ORGANISM Xenopus laevis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus.  
1 (bases 1 to 116)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps@email.nih.gov  
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LNL at: [info@image.llnl.gov](mailto:info@image.llnl.gov)  
Seq primer: -400P from Gibco.

FEATURES  
source  
Location/Qualifiers  
1. 116  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
/clone="IMAGE:4963582"  
/clone\_lib="NICHG XGC Sp1"  
/dev\_stage="adult"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: spleen; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.4 kb. Constructed by Life Technologies."  
36 a 17 c 13 g 50 t

BASE COUNT 157 a 42 c 68 g 114 t  
ORIGIN

Query Match 2.1%; Score 23; DB 10; Length 116;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 934 ataatgtatttattattattat 956  
|||||  
Db 5 ATAAATGATTTATTATTATAT 27  
|||||

RESULT 5  
BI3141427/c  
LOCUS BI3141427 381 bp mRNA linear EST 30-JUL-2001  
DEFINITION 36848 MARCH 2PIG Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION BI3141427  
VERSION BI3141427.1 GI:15034716  
KEYWORDS EST.  
SOURCE pig.  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 381)  
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.  
Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine  
Unpublished (2000)  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: [smith@mail.marc.usda.gov](mailto:smith@mail.marc.usda.gov)  
Single pass sequencing. Bases called and alt\_rimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCCAGTCAGCAGC  
Plate: 108 row: D column: 15  
Seq primer: ATTTAGGTGACACTAGT.  
Location/Qualifiers  
1. 381  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC 2PIG"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."  
157 a 42 c 68 g 114 t

BASE COUNT 157 a 42 c 68 g 114 t  
ORIGIN

Query Match 2.0%; Score 22; DB 10; Length 381;  
Best Local Similarity 100.0%; Pred. No. 12;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 829 ttttttaataaattgtcttttt 850  
 |||||  
 Db 29 TTTTAAATAATGCTTTT 8

RESULT 6  
 BI341431/c  
 LOCUS 431 bp mRNA linear EST 30-JUL-2001  
 DEFINITION 368853 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BI341431  
 VERSION BI341431.1 GI:15034720  
 KEYWORDS EST.  
 SOURCE pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 1 (bases 1 to 431)  
 REFERENCE Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
 AUTHORS Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.  
 and Keele,J.W.  
 TITLE Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@mail.marc.usda.gov  
 Single pass sequencing. Bases called and alt.trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -minmatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCACGACG  
 Plate: 108 row: E column: 14  
 Seq primer: ATTTAGTGTACACTATAG.  
 FEATURES  
 Location/Qualifiers  
 1..431  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARC 2P1G"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI;  
 Library made from pooled tissue from testis, ovary,  
 endometrium, hypothalamus, pituitary, and placenta."  
 BASE COUNT 177 a 51 c 93 g 109 t  
 ORIGIN

Query Match 2.0%; Score 22; DB 10; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 829 ttttttaataaattgtcttttt 850  
 |||||  
 Db 239 TTTTAAATAATGCTTTT 218

RESULT 7  
 A2167394/c  
 LOCUS 446 bp DNA linear GSS 29-AUG-2000  
 DEFINITION SP-0098\_A2\_A12.SP6E Strongylocentrotus purpuratus, purple sea  
 urchin, sperm genomic BAC library Strongylocentrotus purpuratus  
 genomic clone Plate=98 Col=24 Row=A, DNA sequence.  
 ACCESSION A2167394  
 VERSION A2167394.1 GI:8337762  
 KEYWORDS GSS.  
 SOURCE Strongylocentrotus purpuratus.  
 ORGANISM Strongylocentrotus purpuratus

Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 Echinoidea; Euechinoidea; Echinacea; Echinoida;  
 Strongylocentrotidae; Strongylocentrotus.  
 1 (bases 1 to 446)  
 REFERENCE Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,  
 AUTHORS Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray  
 ,G.A., Ettensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and  
 Hood,L.  
 TITLE A sea urchin genome project: Sequence scan, virtual map, and  
 additional resources  
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 97 (17), 9514-9518 (2000)  
 MEDLINE 20402566  
 COMMENT Contact: Cameron, RA, Davidson, EH, Hood, L  
 Division of Biology 156-29  
 California Institute of Technology  
 Pasadena California 91125, USA  
 Tel: (626) 395-8421  
 Fax: (626) 793-3047  
 Email: acameron@caltech.edu  
 Plate: 98 row: A column: 24  
 Seq primer: SP6  
 Class: BAC ends  
 High quality sequence stop: 446.  
 FEATURES  
 Location/Qualifiers  
 1..446  
 /organism="Strongylocentrotus purpuratus"  
 /db\_xref="taxon:7668"  
 /clone="Plate=98 Col=24 Row=A"  
 /clone\_lib="Strongylocentrotus purpuratus, purple sea  
 urchin, sperm genomic BAC library"  
 /note="Organ: sperm; Vector: BAC3.6; BAC Clones in B-Col1  
 DH10B"  
 BASE COUNT 169 a 89 c 92 g 96 t  
 ORIGIN

Query Match 2.0%; Score 22; DB 12; Length 446;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 929 tatttataaattgtctttttat 950  
 |||||  
 Db 143 TATTATAAATGTATTATTAT 122

RESULT 8  
 AQ780852  
 LOCUS 718 bp DNA linear GSS 02-AUG-1999  
 DEFINITION HS\_3113\_B1\_B02\_T7C CIT Approved Human Genomic Sperm Library D Homo  
 sapiens genomic clone plate=3113 Col=3 Row=D, DNA sequence.  
 ACCESSION AQ780852  
 VERSION AQ780852.1 GI:5683812  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 718)  
 REFERENCE Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
 AUTHORS Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
 Hood,L.  
 TITLE Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
 JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)  
 MEDLINE 99380589  
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones may be purchased from Research Genetics (info@resgen.com).



normalization to a Cot = 5. Library constructed by Bêto Soares and M.Fátima Bonaído. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbhH19W."

BASE COUNT 53 a 53 c 51 g 34 t 1 others  
ORIGIN

Query Match 1.9%; Score 21; DB 10; Length 192;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 tcctctcttggccctcttgg 102  
|||||  
Db 59 TCCTCTCTTGGCCCTCTTGG 39  
|||||

RESULT 11  
LOCUS AA341708/c  
DEFINITION EST47365 Fetal muscle Homo sapiens cDNA 5' end similar to similar to myosin light chain 2, mRNA sequence.  
ACCESSION AA341708  
VERSION AA341708.1 GI:1994173  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 210)  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moren-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Snall,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.F., Ferlie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kosak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns  
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)  
MEDLINE 96026280  
COMMENT Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org  
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tcgb/ngi/ngi.html>)  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1. 210  
/organism="Homo sapiens"  
/db\_xref="ATCC (inhost):143553"  
/db\_xref="taxon:9606"  
/clone\_lib="Fetal muscle"  
/tissue\_type="muscle"  
/dev\_stage="fetus, 24 wks"  
/note="vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 55 a 56 c 61 g 34 t 4 others

## ORIGIN

Query Match 1.9%; Score 21; DB 9; Length 210;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 tcctctcttggccctcttgg 102  
|||||  
Db 62 TCCTCTCTTGGCCCTCTTGG 42  
|||||

RESULT 12  
LOCUS AA334932/c  
DEFINITION EST39303 Embryo, 9 week Homo sapiens cDNA 5' end similar to similar to myosin light chain-2, type I, mRNA sequence.  
ACCESSION AA334932  
VERSION AA334932.1 GI:1987250  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 234)  
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moren-Palancas,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Snall,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y., Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dimke,D., Feng,D.F., Ferlie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kosak,D.L., Kunsch,C., Hungjun,J., Li,H., Meisner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and Venter,J.C.

TITLE Initial assessment of human gene diversity and expression patterns  
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)  
MEDLINE 96026280  
COMMENT Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org  
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tcgb/ngi/ngi.html>)  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1. 234  
/organism="Homo sapiens"  
/db\_xref="ATCC (inhost):136696"  
/db\_xref="taxon:9606"  
/clone\_lib="Embryo, 9 week"  
/dev\_stage="embryo, 9 wks"  
/note="vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

## FEATURES

source  
1. 234  
/organism="Homo sapiens"  
/db\_xref="ATCC (inhost):136696"  
/db\_xref="taxon:9606"  
/clone\_lib="Embryo, 9 week"  
/dev\_stage="embryo, 9 wks"  
/note="vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

## FEATURES

source  
1. 234  
/organism="Homo sapiens"  
/db\_xref="ATCC (inhost):136696"  
/db\_xref="taxon:9606"  
/clone\_lib="Embryo, 9 week"  
/dev\_stage="embryo, 9 wks"  
/note="vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 62 a 62 c 69 g 39 t 2 others

## ORIGIN

Query Match 1.9%; Score 21; DB 9; Length 234;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 82 tcctctcttgccctcttgg 102  
 ||||||||||||||||||  
 Db 61 TCCTCTCTTGCCCTCTTGG 41

RESULT 13  
 LOCUS AW486359 240 bp mRNA linear EST 25-APR-2001  
 DEFINITION 72999 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION AW486359  
 VERSION AW486359.1 GI:7056465  
 KEYWORDS EST.  
 SOURCE COW.  
 ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovinae; Bos.

REFERENCE 1 (bases 1 to 240)  
 AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett,  
 G.L., Heaton,M.P., Laegreid,W., Rohrer,G.A., Chitko-McKown,C.G.,  
 Perteau,G., Holt,I., Karameycheva,S., Liang,F., Quackenbush,J. and  
 Keefe,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)  
 MEDLINE 21180013

COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390

Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and trimmed with phred  
 v0.980504 e. Vector identified by cross\_match with the -minscore 20  
 and -minmatch 12 options.

PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCAGCAGC  
 Plate: 33 row: G column: 10  
 Seq primer: ATTAGGTGACACTATAG.

FEATURES  
 source  
 1..240  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="MARC 1BOV"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"

/note="Vector: PCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;  
 Library made from pooled tissue from lymph node, ovary,  
 fat, hypothalamus, and pituitary."  
 BASE COUNT 60 a 39 c 33 g 108 t  
 ORIGIN

Query Match 1.9%; Score 21; DB 9; Length 240;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 934 ataatgtattattattatt 954  
 ||||||||||||||||||  
 Db 30 ATAAATGTATTATTATTATT 50

RESULT 14  
 LOCUS W05048 251 bp mRNA linear EST 23-APR-1996  
 DEFINITION za78g02.r1 Soares\_fetal\_lung\_NbHL19W Homo sapiens cDNA clone  
 IMAGE:298706 5' similar to gb:215030 rnal MYOSIN REGULATORY LIGHT  
 CHAIN 2, VENTRICULAR (HUMAN);, mRNA sequence.  
 ACCESSION W05048  
 VERSION W05048.1 GI:1277978

KEYWORDS  
 SOURCE  
 ORGANISM

EST.  
 human.  
 Homo sapiens

REFERENCE  
 AUTHORS

1 (bases 1 to 251)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman  
 M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,  
 Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasikis,E., Waterston  
 R., Williamson,A., Wohlmann,P. and Wilson,R.

TITLE  
 JOURNAL  
 COMMENT

The WashU-Merck EST Project  
 Unpublished (1995)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: mob.REGA+ET  
 High quality sequence stop: 176.

FEATURES  
 source

1..251  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="GDB:1243630"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:298706"  
 /dev\_stage="19 weeks"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a  
 modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(GT) primer  
 15'-NSTTACCAATCTGAGTGAGCGGCCCAATTTTTTTTTTTT-3',  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT7T3 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaldo. This library was constructed  
 from the same fetus as the fetal heart library, Soares  
 fetal heart NbHL19W."

BASE COUNT 72 a 59 c 70 g 45 t 5 others

Query Match 1.9%; Score 21; DB 10; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 tcctctcttgccctcttgg 102  
 ||||||||||||||||||  
 Db 43 TCCTCTCTTGCCCTCTTGG 23

RESULT 15  
 LOCUS AA331022/c 262 bp mRNA linear EST 21-APR-1997  
 DEFINITION EST34794 Embryo, 6 week I Homo sapiens cDNA 5' end similar to  
 similar to myosin 2, light chain, type I, mRNA sequence.  
 ACCESSION AA331022  
 VERSION AA331022.1 GI:1983264  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 262)  
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult  
 C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White  
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Val,C., Clayton,R.A.,  
 Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald

L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferlie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, F.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl), 3-174 (1995)

96026280

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlavet@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13 Reverse.

Location/Qualifiers

1. .262

/organism="Homo sapiens"

/db\_xref="ATCC (inhost):132654"

/db\_xref="taxon:9606"

/clone\_lib="Embryo, 6 week I"

/dev\_stage="embryo, 6 wks"

/note="Organ: embryo; Vector: pBluescript SK-; Site\_1:

EcoRI; Site\_2: XhoI"

70 a 68 c 73 g 50 t 1 others

BASE COUNT

ORIGIN

Query Match 1.9%; Score 21; DB 9; Length 262;  
Best Local Similarity 100.0%; Pred. No. 39;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 82 tccttctcttgccctcttg 102

|||||

Db 29 TCCTTCTCTTGCCCTCTTGG 9

Search completed: September 23, 2002, 02:07:33  
Job time: 5129 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 23:31:39 ; Search time 2227.14 Seconds  
(without alignments)  
10486.101 Million cell updates/sec

Title: us-09-746-375-1

Perfect score: 1116  
Sequence: 1 tcgagtgtagattgtctga.....ataaacacttgatattcccta 1116

Scoring table:  
IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_btg.\*
- 2: gb\_btg.\*
- 3: gb\_in.\*
- 4: gb\_in.\*
- 5: gb\_ov.\*
- 6: gb\_ov.\*
- 7: gb\_ov.\*
- 8: gb\_ov.\*
- 9: gb\_ov.\*
- 10: gb\_ov.\*
- 11: gb\_ov.\*
- 12: gb\_ov.\*
- 13: gb\_ov.\*
- 14: gb\_ov.\*
- 15: gb\_ov.\*
- 16: gb\_ov.\*
- 17: gb\_ov.\*
- 18: gb\_ov.\*
- 19: gb\_ov.\*
- 20: gb\_ov.\*
- 21: gb\_ov.\*
- 22: gb\_ov.\*
- 23: gb\_ov.\*
- 24: gb\_ov.\*
- 25: gb\_ov.\*
- 26: gb\_ov.\*
- 27: gb\_ov.\*
- 28: gb\_ov.\*
- 29: gb\_ov.\*
- 30: gb\_ov.\*
- 31: gb\_ov.\*
- 32: gb\_ov.\*
- 33: gb\_ov.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1116	100.0	1116	6	AX151713	Sequence
2	1116	100.0	1116	6	AX179578	Sequence
3	1114.4	99.9	1167	9	AF279437	Homo sapi
4	1114	99.8	1152	6	AX092422	Sequence
5	1096.6	98.3	1139	6	AX054620	Sequence
6	1090.4	97.7	1132	6	AX048204	Sequence
7	639	57.3	690	9	HS277247	Homo sapi
8	636.4	57.0	8393	9	AF387519	Homo sapi
9	636.4	57.0	133350	9	AC022511	Homo sapi
10	636.4	57.0	191111	9	AC007458	Homo sapi
11	564	50.5	135146	2	AC087562	Pan trogl
12	525.2	47.1	1111	6	AR165228	Sequence
13	525.2	47.1	1118	10	MMU249492	Mus muscu
14	524.4	47.0	1119	6	AR165226	Sequence
15	524.4	47.0	1121	10	MMU249491	Mus muscu
16	429.2	38.5	778	6	AX179614	Sequence
17	380	34.1	537	6	AX054622	Sequence
18	356.4	31.9	501	6	AX179581	Sequence
19	216.4	19.4	5935	6	AR165234	Sequence
20	216.4	19.4	5935	10	MMU294728	Mus muscu
21	215.6	19.3	7445	6	AR165227	Sequence
22	215.6	19.3	8270	10	MMU294727	Mus muscu
23	207	18.5	5397	9	HS277248	Homo sapi
24	73.2	6.6	194074	9	AC090095	Homo sapi
25	73.2	6.6	198810	9	AC018443	Homo sapi
26	72	6.5	175959	2	AL450429	Homo sapi
27	72	6.5	213056	9	AC103590	Homo sapi
28	72	6.5	213080	30	AL357933	Human DNA
29	71.6	6.4	189662	9	AC015468	Homo sapi
30	71.4	6.4	198405	2	AC104317	Homo sapi
31	71	6.4	5915	11	G88152	CR10003 Dan
32	70.4	6.3	135146	2	AC087562	Pan trogl
33	70.4	6.3	152637	2	AC006734	Caenorhab
34	69.8	6.3	151938	9	HS232384	Human DNA
35	69.8	6.3	199392	2	AC090192	Homo sapi
36	69.6	6.2	150199	9	AC087568	Pan trogl
37	69.4	6.2	4438	3	DDIDP87	Dictyostell
38	69.4	6.2	177164	9	AL365202	Human DNA
39	69.2	6.2	92148	2	AC098159	Rattus no
40	69.2	6.2	200306	2	AC104826	Homo sapi
41	69	6.2	201312	2	AC067900	Homo sapi
42	68.8	6.2	198692	2	AC090330	Homo sapi
43	67.8	6.1	168272	2	AC092451	Homo sapi
44	67.6	6.1	150153	2	AC091869	Homo sapi
45	67.4	6.0	1931	2	AC015021	Drosophila

ALIGNMENTS

RESULT 1	AX151713	AX151713	1116 bp	DNA	Linear	PAT 22-JUN-2001
LOCUS	AX151713	Sequence 14 from Patent WO0140467.				
DEFINITION	AX151713					
ACCESSION	AX151713.1	GI:14533647				
VERSION	AX151713.1					
KEYWORDS	human.					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 1116)					
AUTHORS	Presnell, S.R., Xu, W., Kindsvogel, W. and Chen, Z.					
TITLE	Human cytokine receptor					
JOURNAL	Patent: WO 0140467-A 14 07-JUN-2001;					
FEATURES	ZymoGenetics, Inc. (US)					
source	Location/Qualifiers					
	1..1116					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	21..560					
	/note="unnamed protein product"					
	/codon_start=1					

/protein\_id="CAC42593.1"  
/db\_xref="GI:14533648"  
/translation="MAALQKSVSFLMGTLATSCLLLLALLVOGGAAPISSHCRDLK  
SNFOQPYITNRTFMLEKASLADNNITDVRLEKLFHGVSMSEKYLKQVLFNFTLEE  
VLFPSDRFPQYMOEVPFLARLSNRLSTCHIEGDDLHIQRNVOKLQDVKLGESE  
IKATGELDLLFMSLRNACI"  
BASE COUNT 336 a 229 c 222 g 329 t  
ORIGIN

Query Match 100.0%; Score 1116; DB 6; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 1.7e-211; Indels 0; Gaps 0;  
Matches 1116; Conservative 0; Mismatches 0;

QY 1 tcgagttgagattgtctgcaatgcccgcctctctcttcttggccctcttgggtacagggagagcgtg 60  
DB 1 TCGAGTTAGAAATGCTGCAATGGCCGCCCTTCGACAAATCTGTGAGCTCTTCCCTTATGG 60

QY 61 ggaccctggccaccagctgctcctctcttcttggccctcttgggtacagggagagcgtg 120  
DB 61 GGACCCCTGGCCACCAGCTGCTCCTCTCTTGGCCCTCTTGGTACAGGAGGAGCAGCTG 120

QY 121 cgccatcagctcccatcagcaggttgacaagtcacaactccagcagccctatatcacc 180  
DB 121 CGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAACTTCCAGCAGCCCTATATCACCA 180

QY 181 accgaccttcagctgagcaggttgagcaggttgagcaggttgagcaggttgagcaggt 240  
DB 181 ACCGACCTTCATGCTGAGCAGGTGAGCAGGTGAGCAGGTGAGCAGGTGAGCAGGT 240

QY 241 tcattggggagaaactgtccacggagtcagtaagtgagcgtgctgctgctgagcagc 300  
DB 241 TCATTGGGAGAAACTGTCACCGAGTCAGTATGAGTGCAGTGCAGTGCAGTGCAGTGCAG 300

QY 301 aggtgctgaacttcacccctgaaagagtgctgtccctcaatcagtgaggttccagcctt 360  
DB 301 AGGTGCTGAACCTTCAACCTTCAAGAGGCTGCTTCCCTCAATCTGATAGGTTCACGCTT 360

QY 361 atatgcaggagtggtgctcctctgagcaggtcagcaacagcagcagcagcagcagc 420  
DB 361 ATATGCAGGAGTGCTGCTCCTCTGAGCAGCTCAGCAACAGCAGCTAAGCAGCATGTCTA 420

QY 421 ttgaggtgagtcagctgcatatccagggagtgctgcaaaagctgaaagcagcagcagc 480  
DB 421 TTGAAGGTGATGACCTGATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGCTGAA 480

QY 481 agcttgagagagtgagagagatcaagcaattggagagcagcagcagcagcagcagc 540  
DB 481 AGCTTGGAGAGTGGAGAGATCAAGCAATTTGGAGAACTGGATTTGCTGTTTATGTCTC 540

QY 541 tgagaaatgctgctgcttgcagagcaagcagcagcagcagcagcagcagcagcagc 600  
DB 541 TGAGAAATGCTGCTGCTTGCAGCAGCAAGCTGAAATGAATGAATGAATGAATGAAT 600

QY 601 ctgctagaataaacaattagatgccccaaagcagcagcagcagcagcagcagcagcag 660  
DB 601 CTGCTAGAAATAACAATTAGATGCCCAAGCGATTTTATTTTAAACCAAGAGATGGG 660

QY 661 aagccaaactccatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
DB 661 AAGCCAAACTCCATCATGATGGGTGGATTTCCAAATGAACCCCTGCGTTAGTTACAAAGA 720

QY 721 aaccaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780  
DB 721 AACCAATGCCACTTTGTTTATACACCAAGAGGTAGACTTTCTAAGCATAGATATTTAT 780

QY 781 tgataacatttcattgtaactggtgtcttatcacagagagagagagagagagagagag 840  
DB 781 TGATAACATTTTCATTGTAACCTGCTGCTTCTATACAGAGAAACAATTTATTTTAAATAA 840

QY 841 ttgctcttttccataaaag 900  
DB 841 TTGCTCTTTTCCATAAAAGAGATTAATTTCCATTCCTTTAGGGGAAANAAACCCCTAAATA 900

QY 901 gcttcattgtttccataatcagctacttattattattataaaatgtattattattataaaga 960  
DB 901 GCTTCATGTTTCCATAATCAGTCTTATATTATTATTAATGCTATTATTATTATTATAAGA 960

QY 961 ctgcattttatttattatcattttatttaataatggattttatttataagaacaatcattcgata 1020  
DB 961 CTGCATTTTATTATATCAATTTTATTAAATATGGATTTTATTATAGAACAATCATTCGATA 1020

QY 1021 ttgctacttgagtgtaaggctaatatttgatattatttatgacaataatttatagagctataa 1080  
DB 1021 TTGCTACTTCAGTGTAAAGCTAATATTGATATTATTATGACAATAATTATTAGAGCTATAACA 1080

QY 1081 tgtttattgacctcaataaacaacttgagataccta 1116  
DB 1081 TGTTTATTGACCTCAATAAACACTTGGATATCCTA 1116

RESULT 2  
AX179578 AX179578 1116 bp DNA linear PAT 06-AUG-2001  
LOCUS Sequence 1 from Patent WO0146422.  
DEFINITION AX179578  
ACCESSION AX179578  
VERSION AX179578.1 GI:15132010  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1116)  
AUTHORS Presnell,S.R. and Kindsvogel,W.  
TITLE Cyokline zcyto18  
JOURNAL Patent: WO 0146422-A 1 28-JUN-2001;  
Zymogenetics, Inc. (US)  
FEATURES  
source 1. 1116  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
21.560  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC49973.1"  
/db\_xref="GI:15132011"  
/translation="MAALQKSVSFLMGTLATSCLLLLALLVOGGAAPISSHCRDLK  
SNFOQPYITNRTFMLEKASLADNNITDVRLEKLFHGVSMSEKYLKQVLFNFTLEE  
VLFPSDRFPQYMOEVPFLARLSNRLSTCHIEGDDLHIQRNVOKLQDVKLGESE  
IKATGELDLLFMSLRNACI"  
BASE COUNT 336 a 229 c 222 g 329 t  
ORIGIN

Query Match 100.0%; Score 1116; DB 6; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 1.7e-211;  
Matches 1116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcgagttgagattgtctgcaatgcccgcctctcttcttggccctcttgggtacagggagagcgtg 60  
DB 1 TCGAGTTAGAAATGCTGCAATGGCCGCCCTTCGACAAATCTGTGAGCTCTTCCCTTATGG 60

QY 61 ggaccctggccaccagctgctcctctcttcttggccctcttgggtacagggagagcgtg 120  
DB 61 GGACCCCTGGCCACCAGCTGCTCCTCTCTTGGCCCTCTTGGTACAGGAGGAGCAGCTG 120

QY 121 cgccatcagctcccatcagcaggttgacaagtcacaactccagcagccctatatcacc 180  
DB 121 CGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAACTTCCAGCAGCCCTATATCACCA 180

QY 181 accgaccttcagctgagcaggttgagcaggttgagcaggttgagcaggttgagcaggt 240  
DB 181 ACCGACCTTCATGCTGAGCAGGTGAGCAGGTGAGCAGGTGAGCAGGTGAGCAGGT 240

QY 241 tcattggggagaaactgtccacggagtcagtaagtgagcgtgctgctgctgagcagc 300  
DB 241 TCATTGGGAGAAACTGTCACCGAGTCAGTATGAGTGCAGTGCAGTGCAGTGCAGTGCAG 300

QY 301 aggtgctgaacttcacccctgaaagagtgctgtccctcaatcagtgaggttccagcctt 360  
DB 301 AGGTGCTGAACCTTCAACCTTCAAGAGGCTGCTTCCCTCAATCTGATAGGTTCACGCTT 360

QY 361 atatgcaggagtggtgctcctctgagcaggtcagcaacagcagcagcagcagcagc 420  
DB 361 ATATGCAGGAGTGCTGCTCCTCTGAGCAGCTCAGCAACAGCAGCTAAGCAGCATGTCTA 420

QY 421 ttgaggtgagtcagctgcatatccagggagtgctgcaaaagctgaaagcagcagcagc 480  
DB 421 TTGAAGGTGATGACCTGATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGCTGAA 480

QY 481 agcttgagagagtgagagagatcaagcaattggagagcagcagcagcagcagcagc 540  
DB 481 AGCTTGGAGAGTGGAGAGATCAAGCAATTTGGAGAACTGGATTTGCTGTTTATGTCTC 540

QY 541 tgagaaatgctgctgcttgcagagcaagcagcagcagcagcagcagcagcagcagc 600  
DB 541 TGAGAAATGCTGCTGCTTGCAGCAGCAAGCTGAAATGAATGAATGAATGAATGAAT 600

QY 601 ctgctagaataaacaattagatgccccaaagcagcagcagcagcagcagcagcagcag 660  
DB 601 CTGCTAGAAATAACAATTAGATGCCCAAGCGATTTTATTTTAAACCAAGAGATGGG 660

QY 661 aagccaaactccatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 720  
DB 661 AAGCCAAACTCCATCATGATGGGTGGATTTCCAAATGAACCCCTGCGTTAGTTACAAAGA 720

QY 721 aaccaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 780  
DB 721 AACCAATGCCACTTTGTTTATACACCAAGAGGTAGACTTTCTAAGCATAGATATTTAT 780

QY 781 tgataacatttcattgtaactggtgtcttatcacagagagagagagagagagagagag 840  
DB 781 TGATAACATTTTCATTGTAACCTGCTGCTTCTATACAGAGAAACAATTTATTTTAAATAA 840

QY 841 ttgctcttttccataaaag 900  
DB 841 TTGCTCTTTTCCATAAAAGAGATTAATTTCCATTCCTTTAGGGGAAANAAACCCCTAAATA 900

Db 241 TCATTGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGGCTGCTATCTGATGAAGC 300  
Qy 301 aggtgctgaactccacccttgaaagaagtgctgttccctcaactcatgtatgaggttccagcctt 360  
Db 301 AGGTGCTGAACCTTCAACCTTCAAGAAGTGCTGTCCCTCAATCTGATAGGTTCCAGCCTT 360  
Qy 361 atatgcaggaggtggtgcttcccttgccaggtcagcaacagagctcaagacatgtcata 420  
Db 361 ATATGTCAGAGAGGTGTCCTTCTGGCAGGCTCAGCAACAGGCTAAGACATGTCATA 420  
Qy 421 ttgaaggtgatgaactcatatccagaggaatgtgcaaaaagctgaagacacagtgaaaa 480  
Db 421 TTGAAGGTGATGACTGCATATCAGAGGAATGTGCAAAAGCTGAAGACACAGTGAANA 480  
Qy 481 agcttgagagagtggaagatcaaaagcaattgagagaactggttctgtttatgtctc 540  
Db 481 AGCTTGGAGAGAGTGGAGAGATCAAGCAATTTGGAGAACTGGATTTGCTGTTATGTCTC 540  
Qy 541 tgagaaatgcctgaatttgaccagagcaaaagctgaataaactaaccctttcc 600  
Db 541 TGAGAAATGCCCTGCTATTGACGAGCAAAAGCTGAAAAAATGAATAACTTAACCCCTTTCC 600  
Qy 601 ctgctagaataaacaattagatgcccacaaagcatttttttaaccacaaaggaagatgg 660  
Db 601 CTGCTAGAAATAACAATTAGATGCCCAAGCGATTTTTTTTAAACCAAGAGATGGG 660  
Qy 661 aagcaaaactccatcatgatgggtggattccaaatgaacccctcgcttagtatacaagga 720  
Db 661 AAGCAAACTCCATCATCATGGGTGATTTCCAAATGAACCCCTGCGTTAGTTACAAAGGA 720  
Qy 721 aaccaatgccactttgtttataagaccagaggaaggtagactcttctaagcatgatatatt 780  
Db 721 AACCAATGCCACTTTGTTTATAAGACCAGAGGATAGACTTTCTTAAGCATAGATATTAT 780  
Qy 781 tgaataactttcatgttaactggtgtcttatacacagaaaaaacttttttttaataaa 840  
Db 781 TGATAACATTTCAATGTAACCTGGTGTCTATACAGAGAAACAATTTATTTTAAATAA 840  
Qy 841 ttgtcttttccataaaaaaagattacttccattccctttaggggaaaaaaccccttaata 900  
Db 841 TTGCTTTTTCATAAAAAAGATTACTTTCCATTCCTTTAGGGAAAAAACCCCTAAATAA 900  
Qy 901 gctcatgttccataactcagtaactttattttataatgattattattattataaga 960  
Db 901 GCTTCATGTTTCCATAAATCAGTACTTTATATTAATTAATGATTTATTTATTTATAAGA 960  
Qy 961 ctgcatatttattatcatattttataatggtattttattatagaacacattcgtata 1020  
Db 961 CTGCATTTTATTATATCATTTTATTATATGAGATTTATTTATAGAAACATATTCGATA 1020  
Qy 1021 ttgctacttgagtgtaaggctaatattgatatttatgacaataattatagagctataaca 1080  
Db 1021 TTGCTACTTGAGTGAAGGCTAAATATTCATATTTATGACATAATTTATAGACTATAACA 1080  
Qy 1081 tgtttatttgacctcaataaacacttgatataccta 1116  
Db 1081 TGTTTATTGACCTCAATAAACACTTGGATATCCTA 1116

## RESULT 3

AF279437 1167 bp mRNA linear PRI 09-OCT-2000  
LOCUS AF279437 Homo sapiens interleukin 22 (IL22), complete cds.  
DEFINITION AF279437  
ACCESSION AF279437  
VERSION AF279437.1 GI:10719561  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1167)  
REFERENCE Xie,M.H., Aggarwal,S., Ho,W.H., Foster,J., Zhang,Z., Stinson,J.,  
AUTHORS Wood,W.I., Goddard,A.D. and Gurney,A.L.

TITLE Interleukin (IL)-22, a Novel Human Cytokine That Signals through the Interferon Receptor-related Proteins CRF2-4 and IL-22R  
J. Biol. Chem. 275 (40), 31335-31339 (2000)  
10875937  
REFERENCE 2 (bases 1 to 1167)  
AUTHORS Xie,M.-H., Aggarwal,S., Ho,W.-H., Foster,J., Zhang,Z., Stinson,J., Wood,W.I., Goddard,A.D. and Gurney,A.L.  
TITLE Direct Submission  
JOURNAL Submitted (16-JUN-2000) Molecular Biology, Genentech Inc., 1 DNA Way, South San Francisco, CA 94080, USA

## FEATURES

## source

1..1167

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="12"

/map="12q15"

1..1167

/gene="IL22"

58..597

/gene="IL22"

/note="cytokine"

/codon\_start=1

/product="interleukin 22"

/protein\_id="AAG22064.1"

/db\_xref="GI:10719562"

/translation="MAALQKSVSLFMGLTSLCLLLALLVGGGAAPISSHCLDR

SNFOOPYITNRTFMLEASLADNNTDVLRIKELFHCYSMSERCYLKMKVNLFTLEE

VLRPSDRFQPMQEVVPELARI.SNRLSTCHIEGDDLHIQRNVKLDKTVKLGESGE

IKAGELDLLFMSLRNACI"

BASE COUNT 357 a 244 c 228 g 338 t

## ORIGIN

Query Match 99.9%; Score 1114.4; DB 9; Length 1167;  
Best Local Similarity 99.9%; Pred. No. 3.6e-211;  
Matches 1115; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tcgagttagaattgtctgcaatggccgcccgcgaatactctgtgagctcttcccttatgg 60

Db 38 TCGAGTTAGAATTGCTGCAATGGCCGCCCTGCAGAAATCTGTGAGCTTCTTCCCTTATGG 97

Qy 61 ggaccctggccaccagctgctcctctctctctgtgcccctctgtgtacagggagagcgtg 120

Db 98 GGACCCCTGGCCACCAGCTGCCCTCTCTCTTGGCCCTCTTGGTACAGGAGGAGAGCTG 157

Qy 121 cgcccatcagctcccaactgcaggtctgacaagctccaactccagcagccctatatcaca 180

Db 158 CGCCCATCAGCTCCCACTGTCAGGCTTGACAAGTCCAACTTCACAGAGCCCTATATCA 217

Qy 181 accgcaccttcattgctggctgaaggaggtgcttggctgatacaacaacacagacgttgc 240

Db 218 ACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAAACAACACAGAGTTC 277

Qy 241 tcattggggagaaactgttccacggagtcagtatgagtgagcgtgctatctgtatgaagc 300

Db 278 TCATTGGGGAGAAACTGTTCACGGAGTCAGTATGAGTGAGGCGCTGCTATCTGATGA 337

Qy 301 aggtgctgaactccacctgaagaagtgcttccctcaactcgaatgaggttccagcctt 360

Db 338 AGGTGCTGAACCTTCACTTGAAGAAGTGTCTTCCCTCAATCTGATAGGTTCCAGCCT 397

Qy 361 atatgcaggaggtggtgcttcccttccctggcaggtcctgacaaacaggttaagcatgtc 420

Db 398 ATATGCAGAGGTGGTGGCTTCTGCTGGCAGGCTCAGCAACAGGCTAAGACATGTCAT 457

Qy 421 ttgaaggtgatgaactgcataatccagaggaatgtgcaaaaagctgaagacacagtgaaa 480

Db 458 TTGAAGGTGATGACCTGCATATCCAGAGAAATGTGCAAAAGCTGAAGGACACAGTGAAA 517

Qy 481 agcttgagagagtgagagatcaaaagcaatgtgagagaactggttctgtttatgtctc 540

Db 518 AGCTTGGAGAGTGGAGAGATCAAAAGCAATTTGGAGAACTGGATTGCTGTTTATCTCT 577



```

Db 1118 TGTATTGACCTCAATAAACACTTGGATATCC 1151

RESULT 5
AX054620
LOCUS AX054620 1139 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 1 from Patent WO0070457.
ACCESSION AX054620
VERSION AX054620.1 GI:12228184
KEYWORDS
SOURCE
ORGANISM
unidentified.
unclassified.
REFERENCE
1 (bases 1 to 1139)
AUTHORS Parham,C.L., de waal Malefyt,R. and Marehalli,N.L.
Mammalian interleukin-10 homologs: il-d110 and il-d210
Patent: WO 0073457-A 1 07-DEC-2000;
JOURNAL SCHERING CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..1139
/organism="unidentified"
/db_xref="taxon:32644"
/note="surmised Homo sapiens"
59..598
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC21963.1"
/db_xref="GI:12228185"
/translation="MAALOKSVSFLMGLTATSCLLALLVOGGAAPISSHCRLOK
SNFOQPVITNFTFMLEXSLADNNTDVRLLIGELFVGSNSERCYLKQVNLTFLEE
VLPQSDRFQPMQEVVDFPLARLSNRLSTCHIEGDDLIHQNVOKLDTVKKLGESE
IKAIIGELDLFLFMSLRNACI"
158..595
/mat_peptide
/product="unnamed"
337 a 238 c 227 g 333 t 4 others

BASE COUNT 337 a 238 c 227 g 333 t
ORIGIN

Query Match 98.3%; Score 1096.6; DB 6; Length 1139;
Best Local Similarity 99.5%; Pred. No. 1.2e-207;
Matches 1096; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 tgaagttagaattgctgaatggccgcctgcagaaatctgtgagctcttccctatgg 60
|||||
Db 39 TCGAGTTAGAATTGCTGCAATGGCGCCCTGCAGAAATCTGTGAGCTCTTCCCTATGG 98
|||||

QY 61 ggacctggccaccagctgacctctctcttggccctcttggtacaggaggagcagctg 120
|||||
Db 99 GGACCCCTGGCCACCAGCTGCCCTCTCTTGGCCCTCTTGGTACAGGAGGAGCAGCTG 158
|||||

QY 121 cggccatcagctcccaactgcaggttgacaagtcacaacttccagcagccctatatcaca 180
|||||
Db 159 CGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAACCTCCAGCAGCCCTATATCACA 218
|||||

QY 181 accgcacctcagctgctgctgaaggaggttagcttggtgataaacaacagacgttcgtc 240
|||||
Db 219 ACCGCACCTTCATGCTGGCTGAAGAGGNTAGCTTGGCTGATAAACAACACAGACGTTCTG 278
|||||

QY 241 tcattggggagaacgtttccacgagtcagtatagtgagcgtgctatctatgatgaagc 300
|||||
Db 279 TCATTGGGGAGAACTGTTCACCGAGTCAGTATGAGTGAGCGCTGCTATCTGATGAAGC 338
|||||

QY 301 aggtgtgaacttacccttgaaagagtgctgttccctcaatctgtaggttccagctt 360
|||||
Db 339 AGGTGCTGAACCTTACCCTTGAGAGAGTGTGTTCCCTCAATCTGATAGGTTCCAGCCTT 398
|||||

QY 361 atatgcaggaggtgtgcctcttccctggcaggctcagcaacagcgtgaagcatgtcata 420
|||||
Db 399 ATATGCAGGAGTGTGGCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCATCTCATATA 458
|||||

QY 421 ttgaagggtgtagacctgcataatccagaggaatgtgcaaaagctgaagacagtgaaaa 480
|||||
Db 459 TTGAAGGTGATGACCTGTCATCCAGAGGAATGTGCAAAAGCTGAAGGACAGTGAAAA 518
|||||

```

```

QY 481 agcttgagagagtgagagatcaaaagcaattggagaactggtattgtgtttatgtctc 540
|||||
Db 519 AGCTTGGAGAGAGTGGAGAGATCAAAAGCAATTGGAGAACTGGATTTGCTGTTTATGTCTC 578
|||||

QY 541 tgagaaaatgcctgcatcttgaccagagcaaacgtgaaatgaataactaacccccctttcc 600
|||||
Db 579 TGAGAAATGCCTGCTATTTGACCAGAGCAAGCTGAAAATGAATAACTAACCCCTTTTCC 638
|||||

QY 601 ctgctagaaaataacaatlagatgcccacaaagcgatttttttaacccaaagaaagatggg 660
|||||
Db 639 CTGCTAGAAAATAACAATTAGATGCCCAAGCGATTTTTTAAACCAAAAGAAAGATGGG 698
|||||

QY 661 aagccaaactccatcatgatgggtggattccaaaatgaacccctgcgttagttacaaaaga 720
|||||
Db 699 AAGCCAAACTCCATCATGATGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGR 758
|||||

QY 721 aaccaatgccacttttcttataagaccagaggtagactttcctaagcatagatttat 780
|||||
Db 759 AACCAATGCCACTTTTGTATTATAGACCAGAGGATAGACTTTCTWAGCATAGATATTAT 818
|||||

QY 781 tgataacatttctgaactggtgtctctatcacagaaaacaatttttttttaataaa 840
|||||
Db 819 TGATAACATTTCTGTAACCTGCTTCTATACAGAAAACAATTTATTTTAAATAA 878
|||||

QY 841 ttgtcttttccataaaaaagattactttccattctctttaggggaaaaaacccctaaata 900
|||||
Db 879 TTGTCTTTTCCATAAAAAAGATTACTTTCCATTCTTTAGGGGAAAAAACCCCTAAATA 938
|||||

QY 901 gcttcattgttccataacacgtaactttatattataataatgtattattattataaga 960
|||||
Db 939 GCTTCATGTTTCCATAAATCAGTACTTTATATTATAAATGATTTATTTATTTATAAGA 998
|||||

QY 961 cgcattttatttatcatcttttattataatgattattattataaacaatcttcgata 1020
|||||
Db 999 CTGCATTTTATTATATCATTTTATTATATGATGATTTATTATAGAAACATCATTCGATA 1058
|||||

QY 1021 ttgtcacttgagtgtaaggctaattgtgattttatgacaataattatagagctataaca 1080
|||||
Db 1059 TTGCTACTTGAGTGTAGGCTAATATTGATATTTATGACAATAATTATAGAGCTATAACA 1118
|||||

QY 1081 tgttatttgacctcaataaa 1101
|||||
Db 1119 TGTATTATGACCTCAATGAA 1139
|||||

RESULT 6
AX048204 1132 bp DNA linear PAT 15-DEC-2000
LOCUS AX048204
DEFINITION Sequence 44 from Patent WO0070049.
ACCESSION AX048204
VERSION AX048204.1 GI:11876994
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1132)
AUTHORS Tang,Y.T., Yue,H., Lai,P., Burford,N., Bandman,O., Baughn,M.R.,
Azimzai,Y., Lu,B.A. and Patterson,C.
Extracellular signaling molecules
Patent: WO 0070049-A 44 23-NOV-2000;
Incyte Genomics, Inc. (US)
FEATURES
source
Location/Qualifiers
1..1132
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 5571181CBI"
BASE COUNT 333 a 240 c 226 g 333 t
ORIGIN

Query Match 97.7%; Score 1090.4; DB 6; Length 1132;

```

Best Local Similarity 99.9%; Pred. No. 2.1e-206; Matches 1091; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	tcaggttagaattgtctgcaatgcccgcctcgagaaatctgtgagctcttcccttatgg	60						
Db	31	TCGAGTTAGAAATGTCTGCAATGCCGCCCTCGAGAAATCTGTGAGCTCTTCCTTATGG	90						
Qy	61	ggacctggccaccagctgctctctctcttcttggccctcttggtagagggaggagcagctg	120						
Db	91	GGACCTGGCCACCAGCTGCTCTCTCTTGGCCCTCTTGGTACAGGGAGGAGCAGCTG	150						
Qy	121	cgccatcaagctcccaactgagcgttgaagaagtcacaactccagcagccctatacaca	180						
Db	151	CGCCCATCAGCTCCACTGCAAGGCTTACAAAGTCAACTTCCAGCAGCCCTATATCACCA	210						
Qy	181	acgcaccttcactgctgaagagagctagctggctgatacaacaacacagacgttcgc	240						
Db	211	ACCACCTTCATGCTGGCTAAGAGAGCTAGCTTGGCTGATACACACAGACGCTTCGC	270						
Qy	241	tcattggggagaaactgtccacggagtcagtatgagtgagcgtgtctatctgatgaagc	300						
Db	271	TCATTGGGGAGAACTGTTCACGAGTCAAGTATGAGTGAGCGCTGCTATCTGATGAAGC	330						
Qy	301	aggtgctgaacttcaccccttgaaagaagtgctgttccctcaatctgataggtccagcct	360						
Db	331	AGTGTCTGAACCTTCAACCTTGAAGAAGTGTCTTCCCTCAATCTGATAGGTTCAGCGCTT	390						
Qy	361	atagcaggagtggtggtccctctgcccagctcagcaacagcgttaagcacatgcata	420						
Db	391	ATATGCAAGAGTGTGTCCTCTCCGCCAGGCTCAGCAACAGGCTAAGCACATGTCTATA	450						
Qy	421	ttgaagtgtagactgcatactcacagggaatgtgcataaagctgaaggacagtgaaaa	480						
Db	451	TTGAAGGTGATGACCTGCATATCAGAGAGAAATGTGCAAAAGCTGAAGGACACAGTGA	510						
Qy	481	agcttgagagagtgagagataaagcaattgagaaactgagattgtctgttatgtctc	540						
Db	511	AGCTTGGAGAGAGTGGAGAGATCAAGAGCAATTTGGAGAACTGGATTGTCTTATGTCTC	570						
Qy	541	tgaataatgctgctatttgaccagagcaagctgaaataaactaaactaaacccctttcc	600						
Db	571	TGAGAAATGCTGCTATTGACCAAGCAAGCTGCAAAATGAATAACTAATCCCTCTTCC	630						
Qy	601	ctgctagaaaaatacaattagatgccccaaagcagatttttttaaccaaagaagatagg	660						
Db	631	CTGCTAGAAATAACAATTTAGATGCCCAAGCGATTTTTTTTAAACCAAAAGGAATGGG	690						
Qy	661	aagccaaactccatcatgctgggtgattccaaatgaacccctgcgttagttacaaaaga	720						
Db	691	AAAGCCAAACTCCATCATGATGGGTGGATTCCAATGAACCCCTGCTTGTATTACAAAGA	750						
Qy	721	aaccaatgccacttttggttataagaccagagaggtagactttctaaagcatagatttat	780						
Db	751	AACCAATGCCACTTTGTTTATAAGACCAGAGGTAGACTTCTTAAGCATAGATATTAT	810						
Qy	781	tgatacaatttcattgttaactggtgtcttatcacagaaaaaactttatttttaataa	840						
Db	811	TGATAAATTTTCAATTTGAATGCTTCTATACACAGAAAAACAATTTTATTTTAAATAA	870						
Qy	841	ttgtctttttccataaaaagaattacttcttccattctcttaggggaaaaaacccctaaata	900						
Db	871	TTGTCTTTTTCCTATAAAGAAAGTAACTTTCCTTCTTGGGGAAAAAACCCCTAAATA	930						
Qy	901	gcttcagtgttccataatcagtaactttattattataaaatgtattttattattataaga	960						
Db	931	GCTTCATGTTTCCATAATCAGTACTTTATATTATAAATGATTTATTATTATTATAAGA	990						
Qy	961	ctgcattttatttatcatctttttataatggtattttattatagaacatcatctcgata	1020						
Db	991	CTGCATTTTATTATCATTTTATTATATGATGATTTTATTATAGAAAACATCATTCGATA	1050						
Qy	1021	ttgctacttgagtgtaaggcctaattgattatttatgacataaattatagagctataaca	1080						

Db	1051	TTGCTACTTGAGTGAAGCGCTAATATTGATATTTATGACAATAATTATAGAGCTATAACA	1110						
Qy	1081	tgttattttgac	1092						
Db	1111	TGTTTATTGCC	1122						

RESULT 7

HSA277247

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

1

690 bp

linear

22 (IL-22 gene)

GI:9968293

interleukin 22.

human.

Homo sapiens

1 (bases 1 to 690)

Dumontier, L., Van Roost, E., Colau, D. and Renauld, J.C.

Human interleukin-10-related T cell-derived inducible factor: molecular cloning and functional characterization as an hepatocyte-stimulating factor

Proc. Natl. Acad. Sci. U.S.A. 97 (18), 10144-10149 (2000)

20420346

Renauld, J.C.

2 (bases 1 to 690)

Direct Submission

Submitted (10-APR-2000) Renauld J.C., UCL 74.59, Ludwig Institute for Cancer Research, Avenue Hippocrate, 74, B-1200 Brussels, BELGIUM

Location/Qualifiers

1..690

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/cell\_type="peripheral blood mononuclear cells"

72..611

/gene="IL-22"

72..611

/gene="IL-22"

/note="previously called IL-10-related T-cell-derived inducible factor (IL-TIF)"

/codon\_start=1

/product="interleukin 22"

/protein\_id="CAC06085.1"

/db\_xref="GI:9968293"

/translation="MAALQKSVSFLMGLTATSCLLLLALLVQGAAPAISSHCRLDK SNFOOPYITNRTFMLEAKESLADNNTDVLIGELKLFHGVMSERCYLMKQVLFNLTLEE VLFPSDRFQPYMOEVPVFLARLSNRLSTCHIEGDLHIQRNVQKLKDTVAKLGESGE IKAIGELDLFLMSLRNACI"

182 a

176 c

167 g

165 t

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

0;

1

60

52

61

112

121

172

181

232

1

60

111

120

171

231

240

291



```
QY 241 tcattggggagaacattgtccacggagtcagtgatgagtcgagtcgtctatctgatgaagc 300
Df 292 TCATTGGGAGAAACGTTCCACGGAGTCAGTATGAGTGAGCGCTGCTATCTGATGAAGC 351
QY 301 agtgcgtgaactcacccttgaagaagtcgtctgctccctcaatgataggtccagcctt 360
Df 352 AGTGCCTGAACCTTCACCCCTTGAAGAAGTCGTCTCCCTCAATGTAGTGTCCAGCCTT 411
QY 361 atatgcaggaggtgggtgccccttctgcccaggtcagcaacaggtcaagcacatgtcata 420
Df 412 ATATGCAGGAGGTGGTGGCCCTTCTGCCCAGGCTCAGCACAGCTAAGCACATGTCATA 471
QY 421 ttgaaggtgatgacctgcatactcagaggaatgtcacaagctgaaggacacacagtgaaaa 480
Df 472 TTGAAGGTGATGACCTGTCATATCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAA 531
QY 481 agcttgagagagatgagagatcaagaacattgagaactggattgctgtttatgtctc 540
Df 532 AGCTTGAGAGAGTGAGAGATCAAGACAAATGGAGAACTGGGATTTGCTGTTTATGTCTC 591
QY 541 tgagaatgcctcatttgaccagagcaaaagctgaaaaatgaataaactaaccccccttcc 600
Df 592 TGAGAAATGCTGCTGCTTGGACGAGCAAAAGCTGAAAAATGAATGAATGAATGAATGAAT 651
QY 601 ctgctagaataacaattagatgcccccaagcgattttt 639
Df 652 CTGCTAGAATAACAATTAGATGCCCCAAGCGATTTT 690

RESULT 8
AF387519 AF387519 8393 bp DNA linear PRI_14-JUN-2001
LOCUS Homo sapiens interleukin 22 (IL22) gene, complete cds.
DEFINITION AF387519
ACCESSION AF387519
VERSION AF387519.1 GI:14423570
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 8393)
AUTHORS Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
Yi,Q. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2001) Molecular Biotechnology, University of
Washington, 1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: SeattleSNPs. NHLBI Program for
Genomic Applications, UW-FHCRC, Seattle, WA (URL:
http://pga.mbt.washington.edu).
FEATURES
source 1..8393
/organism="Homo sapiens"
/db_xref="taxon:9606"
repeat_region 1..282
/rpt_family="LIMA7"
/rpt_type="dispersed"
variation 206
/frequency="0.43"
/replace="G"
repeat_region 283..630
/rpt_family="LIMA8"
/rpt_type="dispersed"
variation 301
/frequency="0.29"
/replace="A"
variation 353
/frequency="0.04"
/replace="C"
variation 418
/frequency="0.01"
/replace="T"
variation 523
```

```
/frequency="0.30"
/replace="A"
575
/frequency="0.04"
/replace="T"
717
/frequency="0.47"
/replace="C"
893
/frequency="0.01"
/replace="T"
983
/frequency="0.02"
/replace="A"
997
/frequency="0.12"
/replace="T"
998
/frequency="0.23"
/replace="T"
1047..1144
/rpt_family="LIMCa"
/rpt_type="dispersed"
1121
/frequency="0.03"
/replace="G"
1165
/frequency="0.43"
/replace="A"
1412
/frequency="0.13"
/replace="T"
1500
/frequency="0.01"
/replace="T"
1628
/frequency="0.43"
/replace="T"
1637
/frequency="0.01"
/replace="A"
1749
/frequency="0.01"
/replace="A"
2002
/frequency="0.10"
/replace="G"
join(2056..2298,2732..2797,2914..3057,3983..4048,
6885..6841)
/gene="IL22"
/product="interleukin 22"
2056..6841
/gene="IL22"
join(2113..2298,2732..2797,2914..3057,3983..4048,
6885..6762)
/gene="IL22"
/codon_start=1
/product="interleukin 22"
/protein_id="AAK62468.1"
/db_xref="GI:14423571"
/translation="MAALQKSVSFLMGTLATSCLLALLVGGGAAPISSHCRLDK
SNFQOPYITNRTEFLAKESLADNITDVRILIGKLFHGVSMSERCYLMKOVLPNTLEE
VLFPQSDRFPQYMQEVVFPFLARLSNRLSTCHIIEGDDLHIQRNVQKLKDYTKLGESGE
IKAIQELDLLFMSLRNACI"
2575
/gene="IL22"
/frequency="0.01"
/replace="A"
2576
/gene="IL22"
/frequency="0.02"
/replace="G"
2697
variation
variation
variation
```



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

# REFERENCE AUTHORS

1 (bases 1 to 133350)  
 Alzrydy, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Benton, J., Bimago, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowles, S., Brilev, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, E., Day-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunartne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Huly, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisgied, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Matindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Miner, Z., Moser, M., Neal, D., Nelson, D., Newton, J., Morgan, M., Morris, S., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwono, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoostart, N., Stinson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Zhou, J., Wleczek, R., Wooden, S., Worley, K., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Kucherlapati, R. and Gibbs, R.

Direct Submission  
 2 (bases 1 to 133350)  
 Worley, K.C.

Direct Submission  
 Submitted (05-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 133350)  
 Worley, K.C.

Direct Submission  
 Submitted (11-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jul 11, 2001 this sequence version replaced gi:12656660.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT-----

----- Summary Statistics -----  
 Contig length: 150667  
 Phrap values in estimate: 149705  
 Average error rate (BCM-Phrap estimate): 1.52849e-05  
 Fraction of Phrap values less than 40 : 0.0163588  
 Number of consensus changing edits: 83  
 Number of N's in consensus : 0

----- Consensus changing edits -----  

Position	Original	Context	Edited	Context
7824	tattatat(n)nnngtatact	tattatat(g)tggtatact	tattatat(g)tggtatact	
7825	attatat(n)nnngtatact	attatat(g)tggtatact	attatat(g)tggtatact	
7826	ttatatnn(n)ngtatact	ttatat(g)tggtatact	ttatat(g)tggtatact	
7827	ttatatnn(n)ngtatact	ttatat(g)tggtatact	ttatat(g)tggtatact	
7855	atatatgc(n)acatttata	atatat(g)tggtatact	atatat(g)tggtatact	
8162	agaccattt(n)tgagatgcca	agaccattt(t)tgagatgcca	agaccattt(t)tgagatgcca	
10049	tgggcagatc(n)cnmnnnnnn	tgggcagatc(c)tgagatgcca	tgggcagatc(c)tgagatgcca	
10051	ggcagatcnc(n)nnnnnnnn	ggcagatc(c)tgagatgcca	ggcagatc(c)tgagatgcca	
10052	ggcagatcnc(n)nnnnnnnn	ggcagatc(c)tgagatgcca	ggcagatc(c)tgagatgcca	
10053	cagatcnc(n)nnnnnnnn	cagatc(c)tgagatgcca	cagatc(c)tgagatgcca	
10054	agatcncnn(n)nnnnnnnn	agatc(c)tgagatgcca	agatc(c)tgagatgcca	
10055	gatcncnn(n)nnnnnnnn	gatc(c)tgagatgcca	gatc(c)tgagatgcca	
10057	tcnnnnnn(n)nnnnnnnn	tcnnnnnn(g)tgagatgcca	tcnnnnnn(g)tgagatgcca	
10058	cnnnnnnn(n)nnnnnnnn	cnnnnnnn(g)tgagatgcca	cnnnnnnn(g)tgagatgcca	
10059	cnnnnnnn(n)nnnnnnnn	cnnnnnnn(g)tgagatgcca	cnnnnnnn(g)tgagatgcca	
10060	cnnnnnnn(n)nnnnnnnn	cnnnnnnn(g)tgagatgcca	cnnnnnnn(g)tgagatgcca	
10061	nnnnnnnn(n)nnnnnnnn	nnnnnnnn(g)tgagatgcca	nnnnnnnn(g)tgagatgcca	
10062	nnnnnnnn(n)nnnnnnnn	nnnnnnnn(g)tgagatgcca	nnnnnnnn(g)tgagatgcca	
10064	nnnnnnnn(n)nnnnnnnn	nnnnnnnn(g)tgagatgcca	nnnnnnnn(g)tgagatgcca	
10065	nnnnnnnn(n)nnnnnnnn	nnnnnnnn(g)tgagatgcca	nnnnnnnn(g)tgagatgcca	
10067	nnnnnnnn(n)nnnnnnnn	nnnnnnnn(g)tgagatgcca	nnnnnnnn(g)tgagatgcca	
10068	nnnnnnnn(n)nnnnnnnn	nnnnnnnn(g)tgagatgcca	nnnnnnnn(g)tgagatgcca	
10094	catgggtgaaa(n)ccctgtcta	catgggtgaaa(c)ccctgtcta	catgggtgaaa(c)ccctgtcta	
10102	aaacctgtct(n)tactaaaaa	aaacctgtct(c)tactaaaaa	aaacctgtct(c)tactaaaaa	
10112	ttactaaaaa(n)acaaaaata	ttactaaaaa(t)acaaaaata	ttactaaaaa(t)acaaaaata	
10287	atggaagaga(n)tggtggtata	atggaagaga(t)tggtggtata	atggaagaga(t)tggtggtata	
19854	aaatcgaat(n)aaaaagccca	aaatcgaat(t)aaaaagccca	aaatcgaat(t)aaaaagccca	
19859	tgaaataaaa(n)gccagagtta	tgaaataaaa(t)gccagagtta	tgaaataaaa(t)gccagagtta	
19882	ttctatctc(n)ctgactaaan	ttctatctc(t)ctgactaaan	ttctatctc(t)ctgactaaan	
19889	ctcctgact(n)aanmmnnnn	ctcctgact(t)aanmmnnnn	ctcctgact(t)aanmmnnnn	
19890	ctcctgact(n)aanmmnnnn	ctcctgact(t)aanmmnnnn	ctcctgact(t)aanmmnnnn	
19892	ctcctgact(n)aanmmnnnn	ctcctgact(t)aanmmnnnn	ctcctgact(t)aanmmnnnn	
19893	ctcctgact(n)aanmmnnnn	ctcctgact(t)aanmmnnnn	ctcctgact(t)aanmmnnnn	
19894	tgactnaann(n)nnnnnnnn	tgactnaann(t)nnnnnnnn	tgactnaann(t)nnnnnnnn	
19895	tgactnaann(n)nnnnnnnn	tgactnaann(t)nnnnnnnn	tgactnaann(t)nnnnnnnn	
19896	actnaannnn(n)nnnnnnnn	actnaannnn(t)nnnnnnnn	actnaannnn(t)nnnnnnnn	

19897	ctnaannnn(n)ntnnnnnn	ctagatggcc(a)gttccattag	Db	1935	ATTGATAACATTTTCATTGTAACGTGTTCTTATACACAGAAAACAATTTTATTTTAAAT	1876
19898	tnaannnn(n)tnnnnnnt	tagatggcca(g)ttccattagg	Qy	839	aattgtctttttccataaaagaattacttcttccattccttaggggaaacccctaaa	898
19900	aannnnnn(n)nnnnnnnt	gatggccagt(t)ccattaggcc	Db	1875	AATTGTCTTTTCCATAAAAAGATTACTTTCCATTCCTTTAGGGGAAAACCCCTAAA	1816
19901	nnnnnnnn(n)nnnnntnn	tgccagttc(c)attaggccag	Qy	899	tagttccattttccataatccagctactttatattatataaatgtatttattattattataa	958
19902	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	Db	1815	TAGCTTCATGTTTCCCATATACAGTACTTTTATATATGATTTTATATATATATATA	1756
19903	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	Qy	959	gactgcattttattattattattattattattattattattattattattattattattatt	1018
19904	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	Db	1755	GACTGCATTTTATTTATATCATTTTATTAATATGATTTTATATAGAACATCATTTCCGA	1696
19905	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	Qy	1019	tattgtctacttgagtgcaaggaataattattattattattattattattattattatt	1078
19906	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	Db	1695	TATTGCTACTTGTGAGTGTGAAGCTTAATATTGATATTATGACAAATTAATTATAGACTATA	1636
19907	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	Qy	1079	catgtttatttgacctcaataaacaacacttgatataccta	1116
19908	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	Db	1635	CATGTTTATTGACCTCAATAAACACTTGGATATCCTA	1598
19909	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	RESULT	10		
19910	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AC007458/c			
19911	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	LOCUS			
19912	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	DEFINITION			
19913	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AC007458			
19914	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	VERSION			
19915	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	KEYWORDS			
19916	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	SOURCE			
19917	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	ORGANISM			
19918	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19919	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19920	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19921	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19922	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19923	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19924	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19925	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19926	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19927	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19928	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19929	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19930	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19931	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19932	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19933	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19934	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19935	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19936	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19937	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19938	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19939	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19940	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19941	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19942	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19943	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19944	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19945	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19946	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19947	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19948	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19949	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19950	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19951	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19952	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19953	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19954	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19955	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19956	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19957	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19958	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19959	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19960	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19961	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19962	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19963	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19964	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19965	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19966	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19967	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19968	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19969	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19970	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19971	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19972	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19973	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19974	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19975	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19976	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19977	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19978	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19979	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19980	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19981	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19982	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19983	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19984	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19985	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19986	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19987	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19988	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19989	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19990	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19991	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19992	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19993	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19994	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19995	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19996	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
19997	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			
19998	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	JOURNAL			
19999	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	REFERENCE			
20000	nnnnnnnn(n)nnnnntnn	ggccagttcc(a)ttaggccagg	AUTHORS			

Qy	479	aaagcttgagagagtgagagatcaagcaattggagaactgttgctgtttatgct	538
Db	2235	ACAGCTTGAGAGAGTGAGAGATCAAGCAATTTGGAGACTGGATTGCTTTATGCTC	2176
Qy	539	tctgagaaatcctgcatttaccagagcaagctgaaatgaataactaacccccctt	598
Db	2175	TCTGAGAAATGCTGCATTTGACAGAGCAAGCTGAAATTAATACTAACCCCTTT	2116
Qy	599	ccctgctgagaataacaattagatgcccccaaggcatttttttaaccacaaaggagath	658
Db	2115	CCCTGCTAGAATAACAATTAGATGCCCAAGCCGATTTTAAACCAAGGAAGATG	2056
Qy	659	ggaagccaaactccatcatgatgggtggattccaaatgaaccctgcgttagttacaag	718
Db	2055	GGAAACCAAACTCCATCATGATGGGTGGATTTCCAAATGAACCCCTGCGTTAGTACAAG	1996
Qy	719	gaacccaatgccattttgtttataagaccaggaaggttagacttttcaagcatagattt	778
Db	1995	GAAACCAATGCCACTTTGTTTATAGACCAAGAGTAGACTTTCTTAAGCATAGATATTT	1936
Qy	779	attgataacatttcatttgaactggtgttctatacacagaaacatttttttaaat	838

Query Match 57.08; Score 636.4; DB 9; Length 133350;  
 Best Local Similarity 99.88; Pred. No. 2.6e-116;  
 Matches 637; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

## JOURNAL

Submitted (25-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Jul 31, 1999 this sequence version replaced gi:5230396.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches, of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smil and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT-----

```

----- Summary Statistics -----
Contig length: 151624
Phrap values in estimate: 151315
Average error rate (BCM-Phrap estimate): 0.000165522
Fraction of Phrap values less than 40 : 5
Number of consensus changing edits: 0
Number of N's in consensus : 0

```

```

----- Consensus changing edits -----
Position Original+Context Edited+Context
2273 ggccttcatt(n)tcaattgttt ggccttcatt(t)tcaattgttt
5088 agatgggtgc(n)tactcttccc agatgggtgc(c)tactcttccc
102880 ggtttcactg(n)gttaatacagg ggtttcactg(t)gttaatacagg
123710 tctataatt(n)tggtttatag tctataatt(t)tggtttatag
127972 tttcttgag(n)gaatacctaag tttcttgag(t)gaatacctaag

```

## ----- Distribution of Quality &lt; 40 Bases -----

```

1000|
900|
800|
700|
600|
500|
400|
300|
200|
#
bases

```

```

100|
0|
5 10 15 20 25 30 35 40
phrap Value Range

```

```

----- Version: 1.01 qxfo.
FEATURES Location/Qualifiers
Source 1..191111
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12q15"
/clone="RPC11-444B24"
/complement(1..3489)
/rpt_family="LIP3"
1..82
/repeat_region
misc_feature
/repeat_region
repeat_region complement(3702..4041)
repeat_region /rpt_family="LIP3"
repeat_region complement(4043..7146)
repeat_region /rpt_family="LIPBB"
repeat_region complement(7625..7930)
repeat_region /rpt_family="LIPAI5-16"
repeat_region complement(8427..8844)
repeat_region /rpt_family="LIM2"
repeat_region complement(8860..8988)
repeat_region /rpt_family="MLTID"
repeat_region complement(8997..9284)
repeat_region /rpt_family="LIMC/D"
repeat_region complement(9927..10097)
repeat_region /rpt_family="LIM4"
repeat_region 10575..10852
repeat_region /rpt_family="AluJo"
repeat_region 11022..11060
repeat_region /rpt_family="AT_rich"
repeat_region complement(12158..12355)
repeat_region /rpt_family="MER20"
repeat_region complement(12407..12592)
repeat_region /rpt_family="MIR"
repeat_region 12877..12950
repeat_region /rpt_family="AT_rich"
repeat_region 14057..14169
repeat_region /rpt_family="MIR"
repeat_region 14920..14949
repeat_region /rpt_family="(TG)n"
repeat_region complement(15199..15978)
repeat_region /rpt_family="LIMA9"
repeat_region 16201..16508
repeat_region /rpt_family="AluSx"
repeat_region 16931..17096
repeat_region /rpt_family="MIR"
repeat_region complement(17098..17247)
repeat_region /rpt_family="MIR"
repeat_region 17645..17841
repeat_region /rpt_family="MER107"
repeat_region complement(18996..19099)
repeat_region /rpt_family="MER91B"
repeat_region complement(21141..21217)
repeat_region /rpt_family="L2"
repeat_region 21218..21247
repeat_region /rpt_family="(CA)n"
repeat_region complement(21248..21454)
repeat_region /rpt_family="L2"
repeat_region 21547..21661
repeat_region /rpt_family="LTR9B"
repeat_region complement(21670..22020)
repeat_region /rpt_family="L2"
repeat_region complement(22021..22378)
repeat_region /rpt_family="THE1B"
repeat_region complement(22381..23987)

```

/rpt_family="THEIB-internal"					
complement(23989..24349)					
/rpt_family="THEIB"					
complement(24350..24651)					
/rpt_family="L2"					
25393..25482					
Query Match 57.0% Score 636.4 DB 9 Length 191111;					
Best Local Similarity 99.8% Pred. No. 2.6e-116;					
Matches 637; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	479	aagcttgaggagagtcgaacgcaatggagaactgatttgcgttttatcgc	538		
Db	188245	ACAGCTGCGAGAGTGGAGAGATCAAAAGCATTCGAGAAGCTGGATTGC	188186		
QY	539	tctgagaatgctgcatttgaccagacaagtgaataatgaatacacccttt	598		
Db	188185	TCTGAGAAATGCTGCTTTTGACCAGACAAAGCTGAATAATGAATACTAACCCCTTT	188126		
QY	599	cctcgtagaaatacaaatagatgcccaagcgtattttttaaccaaaaaggagatg	658		
Db	188125	CCTCGTAGAATACAAATAGATGCCCAAGCAGATTTTTTTTAACCAAAGGAGATG	188066		
QY	659	ggaagccaaacctccatcatgatgggtggattccaatgaaccccctgcttagttacaag	718		
Db	188065	GGAAGCCAAACTCCATCATGATGGGTGGATTCCAATGAACCCCCTGTAGTTACAAG	188006		
QY	719	gaaccaaagccacttttttaagaacaggaagtagactttctaagcatagatat	778		
Db	188005	GAACCAATGCCACTTTTCTTTATAGACCAAGGTAGACTTTCTAAGCATAGATATTT	187946		
QY	779	attgataaacatttcattgtaactgggtgctctatcacagaaaaaatattttttaaat	838		
Db	187945	ATTGATAACATTTCAATGTAAGTGGTGTCTATACACAGAAAACATTTATTTTAAAT	187886		
QY	839	aattgtctttttccataaaagattacttccattcccttagggggaaaaacccctaaa	898		
Db	187885	AATTGTCTTTTCCATAAAAAAGATTACTTTTCATTCCTTTAGGGGAAAAAACCCCTAAA	187826		
QY	899	tagcttcatttccataatcagtaacttatattataaagtattattattattataa	958		
Db	187825	TAGCTTCATGTTTCCATAATCAGTACTTTATATTATAAATGATTTATTTATTTATTA	187766		
QY	959	gactgcattttattatcatcttttattgaattgattttattagaacatcttcga	1018		
Db	187765	GATGCATTTTATTATATNCATTTTATTATATGATTTTATTATAGAACATCATTCGA	187706		
QY	1019	tattgctacttgagtgaaggcctaattgatattttatgacataattatagagctataa	1078		
Db	187705	TATTGCTACTTGAGTGAAGGCTAATATTGATATTATGACAATAATTATAGAGCTATAA	187646		
QY	1079	catgtttatttgacctcaataaacacttgagataccta	1116		
Db	187645	CATGTTTTATTGACCTCAATPAAACACTTGGATATCCTA	187608		
RESULT 11					
AC087562					
LOCUS					
DEFINITION					
Pan troglodytes clone RP43-74117, WORKING DRAFT SEQUENCE, 31					
unordered pieces.					
AC087562					
VERSION					
GI:14190648					
KEYWORDS					
HTG; HTGS_PHASE1; HTGS_DRAFT.					
SOURCE					
chimpanzee.					
ORGANISM					
Pan troglodytes					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.					
1 (bases 1 to 135146)					
REFERENCE					
AUTHORS					
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,					
Alisbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbarta,J.,					
Benton,J., Bivage,K., Blankenburg,K., Bonin,D., Bouck,J.,					
Bowling,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,					

/rpt_family="THEIB-internal" complement(23989..24349) /rpt_family="THEIB" complement(24350..24651) /rpt_family="L2" repeat_region 25393..25482					
Query Match 57.0% Score 636.4 DB 9 Length 191111; Best Local Similarity 99.8% Pred. No. 2.6e-116; Matches 637; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
repeat_region					
repeat_region					
repeat_region					
Db 188245 ACAGCTGGAGAGTGGGACATCAAAACAATTGCAGAACTGGATTTCGTGTTATGTC 188186   QY 479 aaagcttgaggagatgagagatacaaaacgaattggagaactgagtgttgcgttttatgctc 538   Db 188245 ACAGCTGGAGAGTGGGACATCAAAACAATTGCAGAACTGGATTTCGTGTTATGTC 188186   QY 539 tctgagaatatgcttcattgaccagagacaagctgaaaaatgaataactaaccccccttt 598   Db 188185 TCTGAGAAATGCTGCTATTTGACCAGACAAAGCTGAATAATGAATACTAACCCCCCTTT 188126   QY 599 cccctgtagaaaatacaaatagatgcccaagcgatttttttaaccaaaagggaagatg 658   Db 188125 CCCTGCTAGCAATACAAATAGATGCCCCCAAGCAGATTTTTTTTAACCAAAGGAAGATG 188066   QY 659 ggaagccaacacctcatgatgggtggattcccaaatgaaccccctgcgttagttcacaaag 718   Db 188065 GGAAGCCAACATCCATCATGATGGTGATTCCTCAATGAACCCCCTGCGTTAGTTACAAG 188006   QY 719 gaaccaaagccactttttttaagacacgaaggtagacttttctaagcatagatatatt 778   Db 188005 GAACCAATGCCACTTTTCTTTAATAGACCAAGAAGTAGACTTTCTAAGCATAGATATTT 187946   QY 779 attgataaacatttcattgtaactgggtgctctatcacagaaaacatttttttttaaat 838   Db 187945 ATTGATAACATTCATGTAAGCTGGTGTTCTATACACAGAAAACATTTATTTTAAAT 187886   QY 839 aattgtctttttccataaaagattacttccattcccttagggggaaaaaacccctcttaa 898   Db 187885 AATTGTCTTTTCCATAAAAAGATTACTTTTCCTTTAGGGGAAAAAACCCCTCTAAA 187826   QY 899 taacttcatttccataatcagtaacttatttataaataatgatttatttattattataa 958   Db 187825 TAGCTTCATGTTTCCATAATCAGTACTTTATATTAATGATTTATTTATTTATATATAA 187766   QY 959 gactgcattttattatcatcttttattgaattgattttattatagaacatcttcga 1018   Db 187765 GATGCATTTTATTATATCATTTTATTAATGATTTTATTTATAGAACAATCTTGA 187706   QY 1019 tattgctacttgagtgaaggcttaattgatattttatgacataattatagagctataa 1078   Db 187705 TATTGCTACTTGAGTGAAGCTAAATTGATTTATGCAATAATTATAGAGCTATAA 187646   QY 1079 catgtttatttgacctcaataaacacttgagataccta 1116   Db 187645 CATGTTTATTGTGCTCAATTAACACTTGGATATCTCA 187608 					
RESULT 11					
AC087562					
LOCUS	Pan troglodytes clone RP43-74117, WORKING DRAFT SEQUENCE, 31	135146 bp	DNA	linear	HTG 19-OCT-2001
DEFINITION	unordered pieces.				
AC087562					
VERSION	AC087562.4 GI:14190648				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	chimpanzee.				
ORGANISM	Pan troglodytes				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.				
AUTHORS	1 (bases 1 to 135146)				
	Muzny,D.N., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alisbrooks,S.L., Amarantunge,H.C., Are,J.R., Banks,T., Barbarta,J., Benton,J., Bivage,K., Blankenburg,K., Bonin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,				

\* be preserved.  
\* 10537: contig of 10537 bp in length  
\* 10538: gap of unknown length  
\* 10539: contig of 5859 bp in length  
\* 10540: contig of 5859 bp in length  
\* 10541: gap of unknown length  
\* 10542: gap of unknown length  
\* 10543: contig of 6039 bp in length  
\* 10544: gap of unknown length  
\* 10545: contig of 6039 bp in length  
\* 10546: gap of unknown length  
\* 10547: contig of 6010 bp in length  
\* 10548: gap of unknown length  
\* 10549: contig of 5239 bp in length  
\* 10550: gap of unknown length  
\* 10551: contig of 4862 bp in length  
\* 10552: gap of unknown length  
\* 10553: contig of 5834 bp in length  
\* 10554: gap of unknown length  
\* 10555: contig of 3606 bp in length  
\* 10556: gap of unknown length  
\* 10557: contig of 5933 bp in length  
\* 10558: gap of unknown length  
\* 10559: contig of 4765 bp in length  
\* 10560: gap of unknown length  
\* 10561: contig of 5560 bp in length  
\* 10562: gap of unknown length  
\* 10563: contig of 4748 bp in length  
\* 10564: gap of unknown length  
\* 10565: contig of 5111 bp in length  
\* 10566: gap of unknown length  
\* 10567: contig of 4381 bp in length  
\* 10568: gap of unknown length  
\* 10569: contig of 4870 bp in length  
\* 10570: gap of unknown length  
\* 10571: contig of 3795 bp in length  
\* 10572: gap of unknown length  
\* 10573: contig of 3792 bp in length  
\* 10574: gap of unknown length  
\* 10575: contig of 3820 bp in length  
\* 10576: gap of unknown length  
\* 10577: contig of 4331 bp in length  
\* 10578: gap of unknown length  
\* 10579: contig of 2940 bp in length  
\* 10580: gap of unknown length  
\* 10581: contig of 3154 bp in length  
\* 10582: gap of unknown length  
\* 10583: contig of 2933 bp in length  
\* 10584: gap of unknown length  
\* 10585: contig of 2682 bp in length  
\* 10586: gap of unknown length  
\* 10587: contig of 2897 bp in length  
\* 10588: gap of unknown length  
\* 10589: contig of 3235 bp in length  
\* 10590: gap of unknown length  
\* 10591: contig of 2848 bp in length  
\* 10592: gap of unknown length  
\* 10593: contig of 2728 bp in length  
\* 10594: gap of unknown length  
\* 10595: contig of 2703 bp in length  
\* 10596: gap of unknown length  
\* 10597: contig of 2617 bp in length  
\* 10598: gap of unknown length  
\* 10599: contig of 2102 bp in length  
\* 10600: gap of unknown length  
\* 10601: contig of 2215 bp in length.  
\* 10602: Location/Qualifiers  
\* 1. .135146  
\* /organism="Pan troglodytes"  
\* /db\_xref="taxon:9598"  
\* /clone="RP43-74117"

BASE COUNT 40919 a 26449 c 25229 g 39519 t 3030 others  
ORIGIN

Query Match 50.5%; Score 564; DB 2; Length 135146;  
Best Local Similarity 98.0%; Pred. No. 6.1e-102;

Matches 592; Conservative 0; Mismatches 10; Indels 2; Gaps 2;  
Qy 514 gagaactgatttgctgttattgtctctgagaaaagcctgcatcttgaccagagaaagt 573  
Dy 28847 GAGAACTGGAATTTGCTGTATATGCTCTGAGAAATGCCTGCATTTGACCAGCAAGCG 28906  
Qy 574 gaaaaatgaataaatacccccttccctgctagataaatacaattagatgccccaaagcg 633  
Dy 28907 GAAAGATGAATAACTAACCCCTTTCTCTGTAGAAATAACAATTAGATGCCCCAAAGCG 28966  
Qy 634 a-ttttttttaacaaaggagatggagagcccaactcccatcatgagtggtggtattcca 692  
Dy 28967 ACTTTTTTTTACCCNAAGGAGATGGGAGCCCAACATCCATCATGATGGTGGATTCCA 29026  
Qy 693 aatgaacccctgcgttagttacaaaaggaaaccaaagcactttttttataagaccagaa 752  
Dy 29027 AATGAACCCCTGTTTTAGTTACAAAGGAACCAATGCCACTTTTCTTTTAAAGACCAGAA 29086  
Qy 753 ggtagacttttaagcatagatattattatgataacatttcattgtaactggtgtctata 812  
Dy 29087 GGTAGACTTTTAAAGCATAGATATTTATTGATAACATTTTCATTTGTAACCTGTTCTATA 29146  
Qy 813 cacagaaacaaatttttttaataaattgtctttttccataaaagagattactttcca 872  
Dy 29147 CACAGAAACAAATTTTAAATAATTTGCTTTTCCATAAAAAGATTACTTTTCCA 29206  
Qy 873 ttcctttaggggaaacccctaaatagcttcatgtttccataaatcagtagctttatatt 932  
Dy 29207 TTCCCTTTAGGGGAAACCCCTAAATAGCTTCATGTTCATATATCAGTACTTTATATT 29266  
Qy 933 tataaatgtattattattattattataagactgctgatttttttttttttttttttttt 992  
Dy 29267 TATAAATGTATTATTATTATTATAAGACTGCTGATTTTATTATATCAATTTTATTATATG 29336  
Qy 993 gatttattatagaacactcattcattatgctactgagtgagtgtaattattgatat 1052  
Dy 29337 GATTTATTATTAGAAGCAATTCGATATGCTACTTCTAGTATAAGGCTTAATATTGATAT 29386  
Qy 1053 ttatgacaataattatagactataacatgttttttttttttttttttttttttttttttt 1112  
Dy 29387 TTATGACAATAATTATAG-GCTATAACATGTTTATTGACCTCAATAAACACTTGGATAT 29445  
Qy 1113 ccta 1116  
Dy 29446 CCTA 29449  
RESULT 12  
ARI65228 ARI65228 1111 bp DNA linear PAT 17-OCT-2001  
LOCUS Sequence 9 from patent US 6274710.  
DEFINITION ARI65228  
ACCESSION ARI65228  
VERSION ARI65228.1 GI:16238721  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1111)  
AUTHORS Dumoutier, L., Louhed, J., and Renauld, J.  
TITLE Antibodies which specifically bind T Cell inducible factors (TIFs)  
JOURNAL Patent: US 6274710-A 9 14-AUG-2001;  
FEATURES Location/Qualifiers  
1. .1111  
/organism="unknown"  
BASE COUNT 346 a 232 c 236 g 297 t  
ORIGIN

Query Match 47.1%; Score 525.2; DB 6; Length 1111;  
Best Local Similarity 73.8%; Pred. No. 2.7e-94;  
Matches 813; Conservative 0; Mismatches 253; Indels 36; Gaps 10;  
Qy 15 tctgcaatggccgcctgcagaaatctgtgagctcttcttcttatgggaccctggccacc 74





```
Db 344 ACCCTGGAAGACATTCTGCTCCGCCAGTCAGACAGGTTCCGGCCCTACATGTCAGGAGGTG 403
Qy 375 gtgccccttctgcccagcgtcagcaacggtgtaagcacatgtcatattgaaagtgatgac 434
Db 404 GTGCCCTTTCTGACCAAACTGCAACATCAGCTCAGCTCCTGTGCATCAGTGGTGACGAC 463
Qy 435 ctgcataccagaggaatgtgcaaaagctgaagcgtgaagcacacagtgaaaaagcttgagagagt 494
Db 464 CAGAACATCCAGAAATGTCAGAGGCTGAAGGAGACAGTCAAGAGCTTGGAGAGAGC 523
Qy 495 ggagagataaagcaatggagaactggatttgcctgtgtttattgtctctgagaaatccctgc 554
Db 524 GGAGAGATCAAAAGCCTCGGGAACCTGACCTGCTGTTATCTCTCTGAGAATGCTTGC 563
Qy 555 atttgaccagagcaagctgaaatgaataactaaccccccttccctgctgagaataac 614
Db 584 GTCTGAGCGAGAAAGCTAGAAACGAAGAACTGCTCTCTCTGCTTCTTAAAGAAAC 643
Qy 615 aattagatgcccaagagatgtttttttaaaccaaaaggaagtggaagccaaactccat 674
Db 644 AATAGATCCCTGAATGGACTTTTTTTA-----CTAAGGAAAGCTGAGAGCTTAACGTCAC 699
Qy 675 catgatgggtgattcccaatgaacccctgctgtagtacaaagaaacccaatgccactt 734
Db 700 CATCATTAAGAGATTTACATGAACCTGGCTCAGTTGAAGAGAAATAGTGTCAA--G 757
Qy 735 ttgttttaagaccagagtgagctttctgaagcatagatatatttattgataacatttcat 794
Db 758 TTGTCCATGAGACCAG-AGGTAGACTTGATAACCAACAAAGATTCATTCACATATTTTAT 816
Qy 795 tgaactgggttctatcacacagaaacaaattttttttaaataattgtctttttccat 854
Db 817 TGTCAATTGAT-----AATGCAACAGAAAAAGTATGACTTTTAAAAAATTGTTT-----G 865
Qy 855 aaaaagattcttccattctttaggggaaaaaaccccttaaatagcttccatgttttcca 914
Db 866 AAGGAGGTTACTCTCATTCCTCTA---GAAGAAAGGCTATGTAACTTCA--TTTCCA 920
Qy 915 taactgactttatttataaattgtatttatttatttataaagactgctatttattta 974
Db 921 TAACCAATACTTTATATATGTAAGTTTATTATTATTATAAGTATA-----CATTTTATTTA 974
Qy 975 tatcaattttataatggattttattatagaaaaacatcattcgatattgctacttgagt 1034
Db 975 TGTCAATTATTAATATGGATTATTATTATAGAAAAATATCTGATGTTGATATTTGAGTA 1034
Qy 1035 taaggctaattgatatttatgacaataattatagagctataacatgtttatttgacct 1094
Db 1035 TAAAGCAATA---ATATTATGATATAAATCATATAG-----AACAAGATATCTTAGGCTT 1087
Qy 1095 caataaacacttggatatacta 1116
Db 1088 TAATAAACACATGAATATCATA 1109

RESULT 14
LOCUS AR165226 1119 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 7 from patent US 6274710.
ACCESSION AR165226
VERSION AR165226.1 GI:16238719
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1119)
AUTHORS Dumoutier,L., Louhed,J. and Renaud,J.
TITLE Antibodies which specifically bind T Cell inducible factors (TIFs)
JOURNAL Patent: US 6274710-A 7 14-AUG-2001;
FEATURES Location/Qualifiers
1..1119
source /organism="unknown"

BASE COUNT 352 a 233 c 232 g 302 t
```

```
ORIGIN
Query Match 47.0%; Score 524.4; DB 6; Length 1119;
Best Local Similarity 74.3%; Pred. No. 3.9e-94;
Matches 819; Conservative 0; Mismatches 246; Indels 37; Gaps 11;

Qy 15 tctgcaatggccgcctcgagaaatctgtgagcttttctctatggggaccctggccacc 74
Db 46 TCTCTGATGGCTGTCTGTCAGAAATCTATAGCTTTTCCCTTTATGGGACTTTGGCGGCC 105
Qy 75 agctccctctctctcttggccctctgtgtacaggagagagcgtgcgccatcagctcc 134
Db 106 AGCTGCTGCTTCTCATTTGCCCTGTGGGCCAGGAGCAAAATGCGCTGCCCTCAACACC 165
Qy 135 cactcagcgttgcaagtccaaactccagcagccctatatcaccaaccgcaccttcag 194
Db 166 CGGTGCAGCTTGAGGTGTCCAACCTCCAGCAGCCGTACATCGTCACCGCACCCTTTATG 225
Qy 195 ctggctaaggagcttagcttggctgataacaacacagacgttctctcatttggggagaa 254
Db 226 CTGGCCAAGGAGGCCAGCCTTCAGATAACAACACAGACGTCGGGCTCATCGGGGAGAA 285
Qy 255 ctgttccaaggagtcagtatgagtgagcgtctatctgataagcagcgtgctgaacttc 314
Db 286 CTGTTCCGAGGAGTCACTGCTAAAGATCAGTCTACCTGATGAAGCAGGTGCTCAACTTC 345
Qy 315 acccttgaagaagtgtcttccctcaactctgatagtttccagccttatatcaggaggtg 374
Db 346 ACCCTGGAAGAGCTGCTGCTCCCGCAGTCAGACAGGTTCCAGCCCTTACATCGAGGAGTG 405
Qy 375 gtgccccttctgcccagcgtcagcaacaggttaagcacatgtcatattgaaagtgatgac 434
Db 406 GTACCTTTCTGACCAAACTCAGCAATCAGCTCAGCTCCTCTGCATCAGCGGTGACGAC 465
Qy 435 ctgcataccagagaaatgtgcaaaagctgaaggacacagtgaaaaagcttggagagact 494
Db 466 CAGAACATCCAGAAATGTCAGAAAGGCTGAAGAGGCTGAAGAGACAGTGAAGAAAGCTTGGAGAGT 525
Qy 495 gtagagatacaagcaatggagaactgtgatttctgtgtttatctctctgagaaatgctgc 554
Db 526 GGAGAGATCAAGCGCATTTGGGAACCTGGACCTGCTGCTTTATGTCCTGAGAATGCTTGC 585
Qy 555 atttgaccagagcaagctgaaaaatgaataactaaaccccccttccctgctgagaaataac 614
Db 586 GTCTGAGCGAGAAAGCTAGAAACGAAGAACTGCTCCTCTCTGCTTCTTAAAGAAAC 645
Qy 615 aattagatgcccaagcagatttttttaaccaaaggaagatgggaaagccaaactccat 674
Db 646 AATAAGATCCCTGAATGGACTTTTTTTA-----CTAAGGAAAGTGAAGAGCTAACGTCCAT 701
Qy 675 catgatgggtgattcccaaatgaacccctgctgtagtacaaaggaacccaatgccactt 734
Db 702 CATCATTAAGATTTTCACTGAACCTGGCTCAGTTGAAAAGGAAAGAAATAGTGTCAA--G 759
Qy 735 ttgttttaagaccagaggtagactttctaaagcagatatttattgataacatttcat 794
Db 760 TTGTCCATGAGACCAG-AGGTAGACTTGATACCCACAAAGATTCATTCACATATTTTAT 818
Qy 795 tgaactgggttctatcacacagaaacaaattttttttaaataaattgctcttttccat 854
Db 819 TGTCACTGATG---ATACACAGAAAAATAATGACTTTTAAAAAATTGTTT-----G 867
Qy 855 aaaaagatttcttccattctttaggggaaaaaacccctaaatagcttccatgttttcca 914
Db 868 AAGGAGGTTACTCTCATTCCTTTA---GAAAAAAGGCTTATGTAACTTCA--TTTCCA 922
Qy 915 taactgactttatttataaagtatttatttatttataaagactgacttttattta 974
Db 923 TATCCAATATTTTATATATGTAAGTTTATTATTATATAAGTATA-----CATTTTATTTA 976
Qy 975 tatcaatttttaatatgatttatttataagaacatcatctcgatatttctgacttgagt 1034
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 23:32:14 ; Search time 217.44 Seconds  
(without alignments)  
8811.978 Million cell updates/sec

Title: US-09-746-375-1

Perfect score: 1116

Sequence: 1 tcgagtagaattgtctgca.....ataaacacttgatattccta 1116

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_012802.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1116	100.0	1116	22	AAD09719 Human cytokine, ZC
2	1116	100.0	1116	22	AAF83741 Human IL-TIF polyp
3	1116	100.0	1177	21	AA081773 Human GIL-13/AE289
4	1114	99.8	1152	22	AAF92134 Human PRO10096 cDN
5	1114	99.8	1152	22	AA087053 Nucleotide sequenc
6	1096.6	98.3	1139	22	AAF28841 Human interleukin-
7	1090.4	97.7	1132	22	AA084310 Human EXCS encodin
8	639	57.3	690	21	AA028839 Human T cell induc
9	639	57.3	690	22	AA014875 Human cDNA encodin

10	526	47.1	1166	21	AAC81774 Murine GIL-19 prot
11	525.2	47.1	1111	21	AAA28817 Murine T cell indu
12	525.2	47.1	1111	22	AAA14860 Mouse cDNA for T c
13	524.4	47.0	1119	21	AAA28815 Murine T cell indu
14	524.4	47.0	1119	22	AAA14858 Mouse cDNA encodin
15	429.2	38.5	778	22	AAD09746 Mouse ZCYTO18 cDNA
16	380	34.1	537	22	AAF28842 Human interleukin-
17	356.4	31.9	501	22	AAD09720 Degenerate sequenc
18	270.4	24.2	504	21	AAA43454 Mouse secreted exp
19	216.4	19.4	5935	21	AAA28818 Murine T cell indu
20	216.4	19.4	5935	22	AAA14878 Mouse partial geno
21	215.6	19.3	7444	21	AAA28816 Murine T cell indu
22	215.6	19.3	7444	22	AAA14859 Mouse partial geno
23	207	18.5	4796	21	AAA28840 Human T cell induc
24	207	18.5	4797	22	AAA14876 Human partial geno
c 25	67.4	6.0	3296	23	ABL17264 Drosophila melanog
c 26	63.2	5.7	6243	20	AAZ09598 Clostridium specie
c 27	62	5.6	6082	24	ABL33752 Human immune syste
c 28	61	5.5	113515	24	ABL34175 Human immune syste
c 29	60.4	5.4	14987	24	ABL32630 Human immune syste
c 30	60.2	5.4	7597	24	ABL33013 Human immune syste
c 31	59.8	5.4	19965	22	AAK73166 Human immune/haema
c 32	59.6	5.3	5413	22	AA046694 Tumour suppressor
c 33	59.6	5.3	20420	22	AAK73165 Human immune/haema
c 34	58.4	5.3	5975	24	AA061096 Human gene regulat
c 35	58.2	5.3	8786	23	ABL16590 Drosophila melanog
c 36	58.2	5.3	740	22	AA042953 DNA encoding G pro
c 37	58.2	5.2	5487	24	ABL33598 Human immune syste
c 38	58.2	5.2	7456	24	ABL33931 Human immune syste
c 39	58	5.2	6641	24	ABL32315 Human immune syste
c 40	57.6	5.2	20420	22	AAK73165 Human immune/haema
c 41	57	5.1	6971	24	ABL33237 Human immune syste
c 42	56.8	5.1	19124	18	AA072882 Plasmodium var-7 g
c 43	56.8	5.1	19124	21	AAZ98287 Plasmodium var-7 p
c 44	56.6	5.1	6175	24	ABL33307 Human immune syste
c 45	56.6	5.1	6298	22	AA043559 Chemically pretrea

#### ALIGNMENTS

RESULT 1  
AAD09719  
ID AAD09719 standard; cDNA; 1116 BP.  
XX AC AAD09719;  
XX DT 10-SEP-2001 (first entry)  
XX DE Human cytokine, ZCYTO18 cDNA.  
XX KW Human; cytostatic; cytokine; ZCYTO18 protein; genetic abnormality;  
KW cancer; inflammation; gene therapy; chromosome 12; ss.  
XX OS Homo sapiens.  
XX FH Key  
XX CDS Location/Qualifiers  
FT 21..560  
FT /\*tag= a  
FT /\*product= "Human ZCYTO18 protein #1"  
FT 21..119  
FT /\*tag= b  
FT /\*mat\_peptide 120..557  
FT /\*tag= c  
FT /\*product= "Human mature ZCYTO18 protein #1"  
FT 57..560  
FT /\*tag= d  
FT /\*product= "Human ZCYTO18 protein #2"  
FT 57..119  
FT /\*tag= e  
FT /\*mat\_peptide 120..557  
FT /\*tag= f  
FT /\*product= "Human mature ZCYTO18 protein #2"



XX (ZYMO ) ZYMOGENETICS INC.

XX Presnell SR, Xu W, Kindsvogel W, Chen Z;

XX WPI: 2001-356158/37.

XX P-PSDB: AAB62664.

XX New soluble cytokine receptor polypeptides and polynucleotides, useful  
PT for diagnosing and treating cancer and inflammatory conditions -

XX Example 17; Page 193-195; 210pp; English.

XX The invention relates to a human cytokine receptor polypeptide,  
CC designated zcytor16. The zcytor16 polypeptide can be expressed by  
CC standard recombinant methodology and can bind to IL-TIF (undefined). The  
CC zcytor16 protein is useful for: inhibiting IL-TIF induced proliferation  
CC or differentiation of hematopoietic cell(s) (progenitors); reducing  
CC IL-TIF induced or IL-9 induced inflammation; and suppressing an  
CC inflammatory response in a mammal with inflammation. Heteromeric/  
CC multimeric receptor polypeptides such as soluble zcytor 16/CRF2-4 can be  
CC used to reduce progression and symptoms of cancer. Zcytor16 polypeptides  
CC can also be used to detect IL-TIF levels which is indicative of  
CC pathological conditions including inflammatory states (e.g. rheumatoid  
CC arthritis) and cancer. Antibodies that bind zcytor16 polypeptides and the  
CC polypeptides themselves are useful for the treatment of inflammation,  
CC inflammatory diseases (e.g. infection, asthma, inflammatory bowel  
CC disease, rheumatoid arthritis and atherosclerosis) and autoimmune  
CC diseases. The antibodies and zcytor16 polynucleotides are also useful  
CC for detecting cancer. The present sequence represents a cDNA encoding  
CC the human IL-TIF protein.

XX Sequence 1116 BP; 336 A; 229 C; 222 G; 329 T; 0 other;

Query Match 100.0%; Score 1116; DB 22; Length 1116;  
Best Local Similarity 100.0%; Pred. No. 2.9e-233; Mismatches 0; Gaps 0;  
Matches 1116; Conservative 0; Indels 0;

QY 1 tcgagttagaattgtctgaatggccgcctgcagaaatctgtgagcttttccctatgg 60  
DB 1 tcgagttagaattgtctgaatggccgcctgcagaaatctgtgagcttttccctatgg 60

QY 61 ggaacctggccaccagctgcctcctctctcttcttggcctcttgtagcaggagagcagctg 120  
DB 61 ggaacctggccaccagctgcctcctctcttcttggcctcttgtagcaggagagcagctg 120

QY 121 cccccatcagctcccaclgcagcttgacaagtccaactccagcagcctatatcaca 180  
DB 121 cccccatcagctcccaclgcagcttgacaagtccaactccagcagcctatatcaca 180

QY 181 accgcaccttcctgctgctgaaggagctgagcttgctgatacaacacagagcttcgtc 240  
DB 181 accgcaccttcctgctgctgaaggagctgagcttgctgatacaacacagagcttcgtc 240

QY 241 tcaattggggagaaactgttccacgagtcagtcagtcagtcagtcagtcagtcagtcagtc 300  
DB 241 tcaattggggagaaactgttccacgagtcagtcagtcagtcagtcagtcagtcagtcagtc 300

QY 301 aggtgctgaacttcaccccttgagaagtcgtgttccctcaactctgataaggttccagcctt 360  
DB 301 aggtgctgaacttcaccccttgagaagtcgtgttccctcaactctgataaggttccagcctt 360

QY 361 atatgcaggaggtggtgcttccttcctggcagctcagcaacaggttaacacatgtcaca 420  
DB 361 atatgcaggaggtggtgcttccttcctggcagctcagcaacaggttaacacatgtcaca 420

QY 421 ttgaagggtgtagctgcataccagaggaatgtgcaaaagctgaagcacacagtgaaaa 480  
DB 421 ttgaagggtgtagctgcataccagaggaatgtgcaaaagctgaagcacacagtgaaaa 480

QY 481 agcttggagagagtgagagatcaagaacttgaggagactgatttgcgttattgtctc 540  
DB 481 agcttggagagagtgagagatcaagaacttgaggagactgatttgcgttattgtctc 540

DB 481 agcttggagagagtgagagatcaagaacttgaggagactgatttgcgttattgtctc 540

QY 541 tgagaaatgctgcgttcatttgaccagagcaagctgaaatgaataactaacccctttcc 600  
DB 541 tgagaaatgctgcgttcatttgaccagagcaagctgaaatgaataactaacccctttcc 600

QY 601 ctgctagaaataacaattagatgcccacgaagcgatttttttaacacaaagaagatggg 660  
DB 601 ctgctagaaataacaattagatgcccacgaagcgatttttttaacacaaagaagatggg 660

QY 661 aagccaaactccatcatgatgggtgattccccaaatgaacccctgctgttagttacaaaaga 720  
DB 661 aagccaaactccatcatgatgggtgattccccaaatgaacccctgctgttagttacaaaaga 720

QY 721 aaccaatgccaactttgtttataagaccagaagtagactttctaaagcatagatttat 780  
DB 721 aaccaatgccaactttgtttataagaccagaagtagactttctaaagcatagatttat 780

QY 781 tgataaacatttcattgtaactggttctctatcacagaaaaacaatttttttaataaa 840  
DB 781 tgataaacatttcattgtaactggttctctatcacagaaaaacaatttttttaataaa 840

QY 841 ttgtcttttccataaaaaagattactttccattccttttaggggaaaaaacccctaaata 900  
DB 841 ttgtcttttccataaaaaagattactttccattccttttaggggaaaaaacccctaaata 900

QY 901 gcttcattgcttcataaactcagctactttattattataaaatgattattattattataaga 960  
DB 901 gcttcattgcttcataaactcagctactttattattataaaatgattattattattataaga 960

QY 961 ctgcattttattattatcatcttttattataatgattattattattataagaacatcctcgata 1020  
DB 961 ctgcattttattattatcatcttttattataatgattattattattataagaacatcctcgata 1020

QY 1021 ttgctacttgagtgtaaggctaaatttgattatttgacataattatagagctataaca 1080  
DB 1021 ttgctacttgagtgtaaggctaaatttgattatttgacataattatagagctataaca 1080

QY 1081 tgtttatttgacctcaataaacaacttggataccta 1116  
DB 1081 tgtttatttgacctcaataaacaacttggataccta 1116

RESULT 3

AAC81773

ID AAC81773 standard; CDNA; 1177 BP.

XX AC AAC81773;

XX DT 23-FEB-2001 (first entry)

XX Human GIL-19/AE289 protein coding sequence.

DE Human; GIL-19/AE289; IL-10; interleukin-10; nutrition;

KW cell proliferation; immune stimulation; immune suppression;

KW haematopoiesis regulation; tissue growth; inflammation; cancer; ss.

OS Homo sapiens.

XX WO200005027-A2.

XX PD 02-NOV-2000.

XX PF 28-APR-2000; 2000WO-US11479.

XX PR 28-APR-1999; 99US-0131473.

XX (GEMY ) GENETICS INST INC.

PA Jacobs K, Fouser L, Spaulding V, Xuan D;

XX WPI: 2000-687325/67.

XX P-PSDB: AAB36292.

XX Human GIL-19 protein that shows a high degree of homology to IL  
PT (interleukin)-10, useful in upregulation of humoral immune responses,  
PT as an antiinflammatory agent and as a modulator of immune responses  
PT associated with injury

XX Claim 1; Page 59; 60pp; English.

XX The present invention provides the protein and coding sequences for the  
CC novel human GIL-19/AE289 protein. The protein shows homology to  
CC interleukin-10 (IL-10) and is assumed to be a cytokine. It can be used  
CC in the regulation of cell proliferation and differentiation,  
CC hematopoiesis, immune stimulation or suppression, tissue growth and  
CC tumour inhibition. In addition, it also has uses in the treatment of  
CC inflammation and in nutrition.

XX Sequence 1177 BP; 362 A; 245 C; 232 G; 338 T; 0 Other;

Query Match 100.0%; Score 1116; DB 21; Length 1177;

Best Local Similarity 100.0%; Pred. No. 3e-233;

Matches 1116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tcgagttagaattgtctgaatggccgcccctgcagaaatctgtgagctcttccctatgg 60  
Db 45 tcgagttagaattgtctgaatggccgcccctgcagaaatctgtgagctcttccctatgg 104  
Qy 61 ggaccctggccaccagctgctccctctctcttggccctcttggtagagggagcagctg 120  
Db 105 ggaccctggccaccagctgctccctctctcttggccctcttggtagagggagcagctg 164  
Qy 121 cggccatcagctcccaactgcaggcttgacaagtcgaactccagcagccctatcacca 180  
Db 165 cggccatcagctcccaactgcaggcttgacaagtcgaactccagcagccctatcacca 224  
Qy 181 accgcaacttcatgtgcttaaggagctagcttggctgataaacaacagacgcttcgtc 240  
Db 225 accgcaacttcatgtgcttaaggagctagcttggctgataaacaacagacgcttcgtc 284  
Qy 241 tcattggggagaaaactgtccacggagtcagatagtgagcgtgctatctgtatgaagc 300  
Db 285 tcattggggagaaaactgtccacggagtcagatagtgagcgtgctatctgtatgaagc 344  
Qy 301 aggtgctgaactccaccttgaagaagtgtgttccctcaatctgataggttccagcctt 360  
Db 345 aggtgctgaactccaccttgaagaagtgtgttccctcaatctgataggttccagcctt 404  
Qy 361 atatcaggaggtgtgtcccttccctggcagcagctcagcaacaggctaaagcatgtcata 420  
Db 405 atatcaggaggtgtgtcccttccctggcagcagctcagcaacaggctaaagcatgtcata 464  
Qy 421 ttgaagggtgatgacctgcataaccagaggaatgtcgaagaagctgaagcacagtgaaaa 480  
Db 465 ttgaagggtgatgacctgcataaccagaggaatgtcgaagaagctgaagcacagtgaaaa 524  
Qy 481 agcttggagagagtgagagatcaagaacattgggagaactgatttgcgtttatgtcttc 540  
Db 525 agcttggagagagtgagagatcaagaacattgggagaactgatttgcgtttatgtcttc 584  
Qy 541 tgagaaatgcttcgctatttgacagagcaagctgaaaatgataactaacccctttcc 600  
Db 585 tgagaaatgcttcgctatttgacagagcaagctgaaaatgataactaacccctttcc 644  
Qy 601 ctgctagaaataacaaattagatgccccaagcgatttttttaacaaaggaagatggg 660  
Db 645 ctgctagaaataacaaattagatgccccaagcgatttttttaacaaaggaagatggg 704  
Qy 661 aagccaaactccatcatgatgggtggattcccaaatgaacccctgcgttagttacaaga 720  
Db 705 aagccaaactccatcatgatgggtggattcccaaatgaacccctgcgttagttacaaga 764  
Qy 721 aaccaatgccactttgttttaaacaccagaaggtagactttcttaagcatagatttat 780

Db 765 aaccaatgccactttgtttataagaccagagtagacttttctaagcatagatatttat 824  
Qy 781 tgataacatttccattgttaactgggttctctatcacagaaaaaactttatttttaataaa 840  
Db 825 tgataacatttccattgttaactgggttctctatcacagaaaaaactttatttttaataaa 884  
Qy 841 ttgtcttttccataaaaaagattactttccattcttccattcttaggggaaaaaacccctcaata 900  
Db 885 ttgtcttttccataaaaaagattactttccattcttccattcttaggggaaaaaacccctcaata 944  
Qy 901 gctcattgttccataaatacagtaactttattataataatgtattattattattataaga 960  
Db 945 gctcattgttccataaatacagtaactttattataataatgtattattattattataaga 1004  
Qy 961 ctgcattttatttatcatcttttataatggattttattatagaacacatcatcgcata 1020  
Db 1005 ctgcattttatttatcatcttttataatggattttattatagaacacatcatcgcata 1064  
Qy 1021 ttgctacttgagtgtaaggctaaattattgataattttgacaataaattatagagctataaca 1080  
Db 1065 ttgctacttgagtgtaaggctaaattattgataattttgacaataaattatagagctataaca 1124  
Qy 1081 tgtttatttgacctcaataaacaacttggatcccta 1116  
Db 1125 tgtttatttgacctcaataaacaacttggatcccta 1160  
RESULT 4  
AAF92134  
ID AAF92134 standard; CDNA; 1152 BP.  
XX  
AC AAF92134;  
XX  
DF 15-MAY-2001 (first entry)  
XX  
DE Human PRO10096 cDNA.  
XX  
KW Human; PRO protein; mapping: ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200116318-A2.  
XX  
PD 08-MAR-2001.  
XX  
PF 24-AUG-2000; 2000WO-US23328.  
XX  
PR 01-SEP-1999; 99WO-US20111.  
PR 15-SEP-1999; 99WO-US21090.  
PR 07-DEC-1999; 99US-0169495.  
PR 09-DEC-1999; 99US-0170262.  
PR 11-JAN-2000; 2000US-0175481.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 03-MAR-2000; 2000US-0187202.  
PR 25-APR-2000; 2000US-0199397.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 05-JUN-2000; 2000US-0209832.  
XX (GETH ) GENENTECH INC.  
XX  
XX Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;  
XX  
XX WPI: 2001-183260/18.  
DR P-PSDB: AAB87602.  
XX  
XX Eighty four nucleic acids encoding PRO polypeptides, useful in  
PT molecular biology, including use as hybridization probes, and in  
PT chromosome and gene mapping.  
XX

PS Claim 2; Fig 153; 278pp; English.

XX The present sequence is the coding sequence for a human PRO polypeptide (secreted and transmembrane). The PRO protein, and PRO agonists, PRO antagonists or anti-PRO antibodies are useful for preparation of a medicament useful in the treatment of a condition which is responsive to the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO protein may also be employed as molecular weight markers for protein electrophoresis. The PRO coding sequence has applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping.

XX Sequence 1152 BP; 342 A; 244 C; 228 G; 338 T; 0 other;

Query Match 99.8%; Score 1114; DB 22; Length 1152;

Best Local Similarity 100.0%; Pred. No. 8.1e-233;

Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcgagttgaattgtctgcaatggcgcgcctgcagaaatctgtgagctttccctatgg 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
38 tcgagttgaattgtctgcaatggcgcgcctgcagaaatctgtgagctttccctatgg 97  
QY 61 ggaacctggccacagctgcctcctctctctgtgaccttctgtacaggagagcagctg 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
98 ggaacctggccacagctgcctcctctctctgtgaccttctgtacaggagagcagctg 157  
QY 121 cgcacctcagctcccaactcagcttgacaagtcacaactccagcagccttatcacca 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
158 cgcacctcagctcccaactcagcttgacaagtcacaactccagcagccttatcacca 217  
QY 181 accgcacctcagctggcctgaaggagctgtgctgatacaacacagcgttcgtc 240  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
218 accgcacctcagctggcctgaaggagctgtgctgatacaacacagcgttcgtc 277  
QY 241 tcattggggagaaactgtccacgagtcagtcattgagtgagcgtctatctgtatgaagc 300  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
278 tcattggggagaaactgtccacgagtcagtcattgagtgagcgtctatctgtatgaagc 337  
QY 301 agtgctgaactccacctgaagagtgctgttccctcaatctgatataggttccagcctt 360  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
338 agtgctgaactccacctgaagagtgctgttccctcaatctgatataggttccagcctt 397  
QY 361 atatcagagagtggtgacctctcctgcccagctcagcaacagcgttaacatgtcata 420  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
398 atatcagagagtggtgacctctcctgcccagctcagcaacagcgttaacatgtcata 457  
QY 421 ttgaaggtgatgaacctgcataatccagagaaatgtgcaaaagctgaaggacagtgaaaa 480  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
458 ttgaaggtgatgaacctgcataatccagagaaatgtgcaaaagctgaaggacagtgaaaa 517  
QY 481 agcttggagagagtgagagatcaaaagcaatggagaaactggtgattgtgtttatgtctc 540  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
518 agcttggagagagtgagagatcaaaagcaatggagaaactggtgattgtgtttatgtctc 577  
QY 541 tgagaaatgcctgcatttgaccagagcaagctgaaaaatgaataactaacccccctttcc 600  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
578 tgagaaatgcctgcatttgaccagagcaagctgaaaaatgaataactaacccccctttcc 637  
QY 601 ctgctagaataaacaattagatgcccacaaagcagatttttttaaccaaagaagatggg 660  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
638 ctgctagaataaacaattagatgcccacaaagcagatttttttaaccaaagaagatggg 697  
QY 661 aagcacaactccatcatgatgggtggattcccaatgaacccctgcgttagttacaaaaga 720  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
698 aagcacaactccatcatgatgggtggattcccaatgaacccctgcgttagttacaaaaga 757  
QY 721 aaccaatgccactttgtttataagaccagaaggttagactttcttaagcatagatatattat 780  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
758 aaccaatgccactttgtttataagaccagaaggttagactttcttaagcatagatatattat 817  
QY 781 tgataaacatttcatgtactggttctctatcacagaaaaaacatttttttaataaa 840  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 818 tgataaacatttcatgttaactggtgtctctatcacagaaaaacaatttttttaataaa 877  
QY 841 ttgtctttttccataaaaaagattactttccattctcttaggggaaaaaacccccctaaata 900  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
878 ttgtctttttccataaaaaagattactttccattctcttaggggaaaaaacccccctaaata 937  
QY 901 gttcattgtttccataaatacagtcactttatattataaaatgtattattattattataaga 960  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
938 gttcattgtttccataaatacagtcactttatattataaaatgtattattattattataaga 997  
QY 961 ctgcattttttattatcaatt 1020  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
998 ctgcattttttattatcaatt 1057  
QY 1021 ttgtctacttgagtgtaaggcctaattgtattgtattttatgacaataattatagagctataaca 1080  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1058 ttgtctacttgagtgtaaggcctaattgtattgtattttatgacaataattatagagctataaca 1117  
QY 1081 tgtttatttgacctcaataaacaacttggatatcc 1114  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1118 tgtttatttgacctcaataaacaacttggatatcc 1151  
RESULT 5  
AAC87053  
ID AAC87053 standard; cDNA; 1152 BP.  
XX  
AC AAC87053;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Nucleotide sequence of human polypeptide PRO10096.  
XX  
KW Human; secreted protein; transmembrane protein; PRO196; PRO444; PRO183;  
PRO185; PRO210; PRO215; PRO217; PRO242; PRO288; PRO365; PRO1361; PRO1308;  
PRO1183; PRO1272; PRO1419; PRO4999; PRO170; PRO248; PRO353; PRO1318;  
PRO1600; PRO9940; PRO533; PRO301; PRO187; PRO337; PRO1411; PRO4356;  
PRO246; PRO265; PRO941; PRO10096; PRO6003; PRO6004; PRO350; PRO3630;  
PRO6309; cell death; genetic disorder; transgenic animal; gene therapy;  
ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 58..597  
FT FT /\*tag= a  
FT sig\_peptide 58..156  
FT FT /\*tag= b  
XX  
PN WO200077037-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 22-MAY-2000; 2000WO-US14042.  
XX  
PR 15-JUN-1999; 99US-0139695.  
PR 20-JUL-1999; 99US-0145070.  
PR 26-JUL-1999; 99US-0145698.  
PR 17-AUG-1999; 99US-0149396.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28301.  
PR 02-DEC-1999; 99WO-US28565.  
PR 07-DEC-1999; 99US-0169495.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05841.



```

PR 20-MAR-2000; 2000MO-US07377.
PR 30-MAR-2000; 2000MO-US08439.
PR 15-MAY-2000; 2000MO-US13358.
PR 17-MAY-2000; 2000MO-US13705.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein DA, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Gurney AL, Kijavini J, Mather JP, Napier MA, Pan J;
PI Pooni NF, Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM;
PI Wood WI, Zhang Z;
XX
XX WPI; 2001-050091/06.
DR P-PSDB; AAB31210.
XX
XX Isolated nucleic acid molecule encoding a PRO polypeptide which is a
PT transmembrane polypeptide is useful for gene therapy and identification
PT of related polypeptides .
XX
XX Claim 2; Fig 63; 244pp; English.
XX
XX The present sequence encodes a human secreted and transmembrane
CC polypeptide. The specification describes human polypeptides, designated
CC PRO196, PRO444, PRO183, PRO210, PRO215, PRO217, PRO242, PRO288,
CC PRO365, PRO1361, PRO1308, PRO1183, PRO1272, PRO1419, PRO4999, PRO170,
CC PRO248, PRO353, PRO1318, PRO1600, PRO9940, PRO533, PRO301, PRO187,
CC PRO337, PRO1411, PRO4356, PRO246, PRO265, PRO941, PRO10096, PRO6003,
CC PRO6004, PRO350, PRO2630 and PRO6309. The biological activity of cells
CC can be modulated with agents that bind to these polypeptides, resulting
CC in the death of the cells. The polynucleotides encoding of these
CC polypeptides are useful in the recombinant production of the
CC polypeptides, as a hybridisation probe to screen libraries to isolate
CC homologous sequences, or to map the gene. They may also be used for
CC analysing genetic disorders, and to produce transgenic animals which are
CC useful for the development and screening of therapeutically useful
CC reagents. The polynucleotides can also be used in gene therapy e.g. to
CC replace a defective gene.
XX
XX Sequence 1152 BP; 342 A; 244 C; 228 G; 338 T; 0 other;
XX
Query Match 99.8%; Score 1114; DB 22; Length 1152;
Best Local Similarity 100.0%; Pred. No. 8.1e-233;
Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 tcgagttagaattgtctgaatggccgctgcagagaatctgtgagctcttccttatgg 60
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 38 tcgagttagaattgtctgaatggccgctgcagagaatctgtgagctcttccttatgg 97
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 61 ggacctggccaccagctgctctctcttggccctcttggtacagggaggagcagctg 120
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 98 ggacctggccaccagctgctctctcttggccctcttggtacagggaggagcagctg 157
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 121 cggccatcagctccacagctgagcttgacaaagtcacaaacttcacagagccctatcacca 180
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 158 cggccatcagctccacagctgagcttgacaaagtcacaaacttcacagagccctatcacca 217
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 181 accgacattcatgctgctgaaggagctgcttgctgctgaatacaacacagagcttcgctc 240
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 218 accgacattcatgctgctgaaggagctgcttgctgctgaatacaacacagagcttcgctc 277
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 241 tcattggggagaaactgttccacgagctgagctgagtgagcgctgctatctgtagaagc 300
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 278 tcattggggagaaactgttccacgagctgagctgagtgagcgctgctatctgtagaagc 337
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 301 aggtgtaacttcaccttgaagaagtgctgttccctcaatctgataggttccagcctt 360
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 338 aggtgtaacttcaccttgaagaagtgctgttccctcaatctgataggttccagcctt 397
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 361 atatcagagagtggtgcttcctcctggccagctcagcaacagggctaaacacatgtcata 420
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy 398 atatcagagagtggtgcttcctcctggccagctcagcaacagggctaaacacatgtcata 457
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||

```

## RESULT 6

AAF28841

ID AAF28841 standard; DNA; 1139 BP.

XX

AC AAF28841;

XX

DT 25-APR-2001 (first entry)

XX

DE Human interleukin-10 homologue IL-D110 DNA sequence.

XX

KW Anti-inflammatory; immunosuppressive; cytostatic; antirheumatic; human;  
 KW antiarthritic; dermatological; thyromimetic; gene therapy; inflammation;  
 KW interleukin 10 homologue; IL-D110; IL-D210; cytokine; tumour therapy;  
 KW haematopoietic cell; probe; immunogen; systemic lupus erythematosus;  
 KW rheumatoid arthritis; Hashimoto's autoimmune response; ss.

OS Homo sapiens.

XX

XX Key Location/Qualifiers

FH CDS 59..598

FT /\*tag= a

FT /\*product= \*IL-D110\*

FT /\*transl\_except= (pos:245..247,aa:Xaa)



FT sig\_peptide /note= "Xaa=unknown"  
 FT 59..157 /\*tag= b  
 FT mat\_peptide 158..595 /\*tag= c  
 XX

PN WO200073457-A1.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14729.

XX 27-MAY-1999; 99US-0322806.

XX (SCHE ) SCHERING CORP.

XX Parham CL, De Waal Malefyt R, Marehalli NL;  
 XX WPI; 2001-061539/07.

DR P-PSDB; AAB37122.

XX New interleukin-D110 genes and polypeptides useful for treating  
 PT conditions associated with abnormal physiology, particularly e.g.  
 PT inflammatory or autoimmune disorders

XX Claim 3; Page 51-52; 62pp; English.

XX This sequence represents the gene sequence for a human interleukin 10  
 CC homologue IL-D110. The invention also relates to another IL-10 homologue  
 CC ID-210 (AAF2842). The IL-D110 is useful for treating conditions  
 CC associated with abnormal physiology or development, including  
 CC inflammatory conditions, either acute or chronic. The new cytokine  
 CC also plays a role in the regulation or development of haematopoietic  
 CC cells, e.g. lymphoid or myeloid cells, which affect immunological  
 CC responses, such as inflammation and/or autoimmune disorders. These may  
 CC also be used in drug screening to identify compounds having binding  
 CC affinity to or other relevant biological effects on IL-D110 functions,  
 CC in anti-tumour therapy, as probes for detecting IL-D110 levels in samples  
 CC from patients suspected of having an abnormal condition, e.g. autoimmune  
 CC or inflammatory, in raising or screen antibodies, as immunogen, in  
 CC diagnostic assays, and in detecting, isolating or identifying a DNA clone  
 CC encoding IL-D110 or IL-D210 from a natural source. IL-D110 antagonists  
 CC may be used to block immune responses, e.g. in situations as inflammatory  
 CC or autoimmune responses, including rheumatoid arthritis, systemic lupus  
 CC erythematosus, or Hashimoto's autoimmune responses.

XX Sequence 1139 BP; 337 A; 238 C; 227 G; 333 T; 4 other;

Query Match 98.3%; Score 1096.6; DB 22; Length 1139;  
 Best Local Similarity 99.5%; Pred. NO. 4.9e-229;  
 Matches 1096; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 tcgagtgagaattgtctgaatggccgcctgcagaaatctgtgagcctttcttattgg 60

DB 39 tcgagtgagaattgtctgaatggccgcctgcagaaatctgtgagcctttcttattgg 98

QY 61 ggacctggccaccagctgctctcttcttggcccttgggtacagggagagcagctg 120

DB 99 ggacctggccaccagctgctctcttcttggcccttgggtacagggagagcagctg 158

QY 121 cgcctatagctcccaactgcagcttgacaagtcacaactccagcagcctatatacaca 180

DB 159 cgcctatagctcccaactgcagcttgacaagtcacaactccagcagcctatatacaca 218

QY 181 accgcaccttcctgctgctgagggcctgctgctgctgataacacacagagcttcgc 240

DB 219 accgcaccttcctgctgctgaggggctgctgctgataacacacagagcttcgc 278

QY 241 tcaatggggagaaactgttccacagagtcagatagtgagcgtgctgctatcgatgaagc 300

DB 279 tcaatggggagaaactgttccacagagtcagatagtgagcgtgctgctatcgatgaagc 338

QY 301 aggtgctgaacttcacccttgaagaagtgtgttccctcaatctgataggtgtccagcctt 360  
 DB 339 aggtgctgaacttcacccttgaagaagtgtgttccctcaatctgataggtgtccagcctt 398  
 QY 361 atatcagagagtggtgccccttccctgcccagctcagcaaacaggttaagcaatgtcata 420  
 DB 399 atatcagagagtggtgccccttccctgcccagctcagcaaacaggttaagcaatgtcata 458  
 QY 421 ttgaaggtgatgacctgcataatccagaggaatgtgcataaagtctgaagggacacagtgaaaa 480  
 DB 459 ttgaaggtgatgacctgcataatccagaggaatgtgcataaagtctgaagggacacagtgaaaa 518  
 QY 481 agcttgagagagtgagagatcaaaagcaatggagaaactggtattgctgtttatgtctc 540  
 DB 519 agcttgagagagtgagagatcaaaagcaatggagaaactggtattgctgtttatgtctc 578  
 QY 541 tgagaaatgctgcatttgaccagagcaaaagtgaataaactaactaacccctttccc 600  
 DB 579 tgagaaatgctgcatttgaccagagcaaaagtgaataaactaactaacccctttccc 638  
 QY 601 ctgctagaaaaataacaattagatgcccccaaaagcgatttttttaacccaaaaggaagatggg 660  
 DB 639 ctgctagaaaaataacaattagatgcccccaaaagcgatttttttaacccaaaaggaagatggg 698  
 QY 661 aagccaaactccatcatgatgggtgattccaaaatgaacccctgctgtattacaaagga 720  
 DB 699 aagccaaactccatcatgatgggtgattccaaaatgaacccctgctgtattacaaagga 758  
 QY 721 aaccaatgccaattttgtttataagaccagaaggttagactttcttaagcattagatttat 780  
 DB 759 aaccaatgccaattttgtttataagaccagaaggttagactttcttaagcattagatttat 818  
 QY 781 tgatacaatttcattgtaactggtgtctctacacagaaaaacaatttttttaataaa 840  
 DB 819 tgatacaatttcattgtaactggtgtctctacacagaaaaacaatttttttaataaa 878  
 QY 841 tggcttttttccataaaaaagattactttccattctcttaggggaaaaaacccctaaata 900  
 DB 879 tggcttttttccataaaaaagattactttccattctcttaggggaaaaaacccctaaata 938  
 QY 901 gcttcagtttccataaactcagctatttattataaattttattattattattattataaga 960  
 DB 939 gcttcagtttccataaactcagctatttattataaattttattattattattattataaga 998  
 QY 961 ctgcattttattattattattattattattattattattattattattattattattattataaga 1020  
 DB 999 ctgcattttattattattattattattattattattattattattattattattattattataaga 1058  
 QY 1021 ttgctacttgagtgtaaggcctaattatgtattattatgacaataattatagactataaca 1080  
 DB 1059 ttgctacttgagtgtaaggcctaattatgtattattatgacaataattatagactataaca 1118  
 QY 1081 tgtttatttgacctcaataaa 1101  
 DB 1119 tgtttatttgacctcaataaa 1139

RESULT 7

AAC84310

ID AAC84310 standard; cDNA; 1132 BP.

XX AAC84310;

XX 19-MAR-2001 (first entry)

XX Human EXCS encoding cDNA (clone ID 5571181CB1).

XX Extracellular signaling molecule; EXCS; anti-inflammatory; human;  
 KW immunosuppressive; cytostatic; neuroprotective; gastrointestinal;  
 KW virucide; antibacterial; anti-HIV; human immunodeficiency virus;  
 KW antifertility; cerebroprotective; nootropic; antitumor; antifungal;  
 KW anticonvulsant; tranquilizer; neuroleptic; vasotropic; gynecological;  
 KW keratolytic; protozoacide; gene therapy; ss.





PA (RENAUD) RENAULD J.  
XX Dumoutier L, Louahed J, Renaud J;  
XX  
XX  
XX WPI: 2001-638496/73.  
DR P-PSDB; AAU09091.  
XX  
XX New isolated nucleic acid molecules encoding T cell inducible factors,  
PT useful as markers for expression or effect of interleukin (IL)-9 in a  
PT subject and diagnosing susceptibility to asthma or allergy  
XX  
XX Claim 1: Page 17; 26pp; English.  
PS  
XX The invention relates to an isolated nucleic acid molecule, which encodes  
CC a T cell derived inducible factor (TIF) which are upregulated by the  
CC cytokine interleukin-9 (IL-9) and induce SPAT transcription factor  
CC activation. The TIF proteins (or their mutants) may be used to test IL-9  
CC ant/agonists for their potency against lymphomas, immune system  
CC disorders, allergies, asthma, acquired immunodeficiency syndrome (AIDS),  
CC autoimmune diabetes and thyroiditis. TIF molecules promote regeneration  
CC or inhibit differentiation of tissue types in which they are active and  
CC therefore be used to develop treatments for melanomas and hepatomas.  
CC The present sequence encodes human TIF.  
XX  
XX Sequence 690 BP; 182 A; 176 C; 167 G; 165 T; 0 other;

Query Match 57.3%; Score 639; DB 22; Length 690;  
Best Local Similarity 100.0%; Pred. No. 1.3e-129; Indels 0; Gaps 0;  
Matches 639; Conservative 0; Mismatches 0;  
Oy 1 tcgagttagaattgtctgcaatggccgctgcagaaatctgtgagctcttccctatgg 60  
Db 52 tcgagttagaattgtctgcaatggccgctgcagaaatctgtgagctcttccctatgg 111  
Oy 61 ggaacctggccacagctgcctctctctctgtgacctctgtgtacaggagagcagctg 120  
Db 112 ggaacctggccacagctgcctctctctctgtgacctctgtgtacaggagagcagctg 171  
Oy 121 cgcctcatcagctcccaactgcagagctgtgcaagctgcaactccagcagccctatatcaca 180  
Db 172 cgcctcatcagctcccaactgcagagctgtgcaagctgcaactccagcagccctatatcaca 231  
Oy 181 accgacacctcatgctggctgaaggagctgctgtgctgatacaacacagacgttgcgc 240  
Db 232 accgacacctcatgctggctgaaggagctgctgtgctgatacaacacagacgttgcgc 291  
Oy 241 tcattggggagagaactgttccacgagtcagtcagtcagtcagtcagtcagtcagtcagtc 300  
Db 292 tcattggggagagaactgttccacgagtcagtcagtcagtcagtcagtcagtcagtcagtc 351  
Oy 301 aggtgtgaacttcaacttgaagagtcgtgttccctcaactctgatagtgttccagcctt 360  
Db 352 aggtgtgaacttcaacttgaagagtcgtgttccctcaactctgatagtgttccagcctt 411  
Oy 361 atatcagagagtggtgcttcccttccctgagagctcagcaacaggttaagcaatgtcata 420  
Db 412 atatcagagagtggtgcttcccttccctgagagctcagcaacaggttaagcaatgtcata 471  
Oy 421 ttgaagggtgagtcgctgataatccagaggaatgtgcataaagctgaaggacagtgaaa 480  
Db 472 ttgaagggtgagtcgctgataatccagaggaatgtgcataaagctgaaggacagtgaaa 531  
Oy 481 agcttggagagagtgagagatcaagcaatgtggagaactggtattgtgtttatgtcttc 540  
Db 532 agcttggagagagtgagagatcaagcaatgtggagaactggtattgtgtttatgtcttc 591  
Oy 541 tgagaaatcctgcattttgaccagagcaagctgaaatgaataactaacccctttccc 600  
Db 592 tgagaaatcctgcattttgaccagagcaagctgaaatgaataactaacccctttccc 651  
Oy 601 ctgctagaataacaaattagatgcccaaacgattttt 639

Db 652 ctgctagaataacaattagatgcccacaaagcgtattttt 690

RESULT 10  
AAC8177A  
ID AAC81774 standard; cDNA; 1166 BP.  
XX  
XX AAC81774;  
AC  
XX 23-FEB-2001 (first entry)  
DT  
XX Murine GIL-19 protein coding sequence.  
DE  
XX Mouse; GIL-19/AE289; IL-10; Interleukin-10; nutrition;  
KW cell proliferation; immune stimulation; immune suppression;  
KW haematopoiesis regulation; tissue growth; inflammation; cancer; ss.  
XX Mus sp.  
OS  
XX WO200065027-A2.  
FN  
XX 02-NOV-2000.  
PD  
XX 28-APR-2000; 2000WO-US11479.  
PF  
XX 28-APR-1999; 99US-0131473.  
PR  
XX (GEMY) GENETICS INST INC.  
PA  
XX Jacobs K, Fouser L, Spaulding V, Xuan D;  
PI  
XX WPI: 2000-687325/67.  
DR  
XX Human GIL-19 protein that shows a high degree of homology to IL  
PT (interleukin)-10, useful in upregulation of humoral immune responses,  
PT as an antiinflammatory agent and as a modulator of immune responses,  
PT associated with injury -  
XX  
XX Disclosure: Fig 1: 60pp; English.  
PS  
XX The present invention provides the protein and coding sequences for the  
CC novel human GIL-19/AE289 protein. The protein shows homology to  
CC interleukin-10 (IL-10) and is assumed to be a cytokine. It can be used  
CC in the regulation of cell proliferation and differentiation,  
CC haematopoiesis, immune stimulation or suppression, tissue growth and  
CC tumour inhibition. In addition, it also has uses in the treatment of  
CC inflammation and in nutrition.  
XX  
XX Sequence 1166 BP; 375 A; 241 C; 244 G; 306 T; 0 other;

Query Match 47.1%; Score 526; DB 21; Length 1166;  
Best Local Similarity 74.4%; Pred. No. 5.7e-105;  
Matches 820; Conservative 0; Mismatches 245; Indels 37; Gaps 11;  
Oy 15 tctgcaatggccgctgcagaaatctgtgagcttcttcttcttcttcttcttcttcttcttcttct 74  
Db 69 tctgctgagctgtcctgcagaaatctgtgagcttcttcttcttcttcttcttcttcttcttct 128  
Oy 75 agctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 134  
Db 129 agctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 188  
Oy 135 cactgagagcttgacaagcttccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 194  
Db 189 cgggtcaagcttgagggtgtcccaacttccagcagcagcagcagcagcagcagcagcagcagc 248  
Oy 195 ctggcttaagagagctgagcttgctgatacaacacagagcttcttcttcttcttcttcttcttct 254  
Db 249 ctggcgaagagagcagccttgcagatacaacacagagcttcttcttcttcttcttcttcttct 308  
Oy 255 ctgttccacagagctgagctgagctgcttcttcttcttcttcttcttcttcttcttcttcttct 314















Mon Sep 23 09:43:54 2002

Db 697 catcattagaagatttcacatgaacctggctcagttgaaaaagaaa 744

Search completed: September 23, 2002, 00:43:53  
Job time: 4299 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	639	57.3	690	4	US-09-354-243B-24	Sequence 24, Appl
2	525.2	47.1	1111	4	US-09-178-973B-9	Sequence 9, Appl
3	525.2	47.1	1111	4	US-09-354-243B-9	Sequence 9, Appl
4	524.4	47.0	1119	4	US-09-178-973B-7	Sequence 7, Appl
5	524.4	47.0	1119	4	US-09-354-243B-7	Sequence 7, Appl
6	216.4	19.4	5935	4	US-09-178-973B-17	Sequence 17, Appl
7	216.4	19.4	5935	4	US-09-354-243B-17	Sequence 17, Appl
8	215.6	19.3	7445	4	US-09-178-973B-8	Sequence 8, Appl
9	215.6	19.3	7445	4	US-09-354-243B-8	Sequence 8, Appl
10	207	18.5	4797	4	US-09-354-243B-8	Sequence 8, Appl
11	63.2	5.7	6243	2	US-09-056-075-1	Sequence 1, Appl
12	56.8	5.1	19124	2	US-08-487-826B-13	Sequence 13, Appl
13	51.8	4.6	4254	2	US-08-443-639-7	Sequence 7, Appl
14	51.8	4.6	8920	2	US-08-446-855A-1	Sequence 1, Appl
15	51.8	4.6	8920	2	US-09-150-741-1	Sequence 1, Appl
16	50.2	4.5	767	4	US-08-998-416-472	Sequence 472, App
17	49	4.4	834	4	US-08-998-416-305	Sequence 305, App
18	47.6	4.3	1368	3	US-08-874-563-5	Sequence 5, Appl
19	47.6	4.3	1368	3	US-08-577-483-14	Sequence 14, Appl
20	47.4	4.2	615	4	US-08-998-416-186	Sequence 186, App
21	47.4	4.2	636	4	US-08-998-416-1137	Sequence 1137, Ap
22	47.4	4.2	701	4	US-08-998-416-701	Sequence 701, App
23	47.4	4.2	837	4	US-08-998-416-288	Sequence 288, App
24	46	4.1	782	4	US-09-007-119-15	Sequence 15, Appl
25	45.6	4.1	615	4	US-08-998-416-186	Sequence 186, App
26	45.6	4.1	636	4	US-08-998-416-1137	Sequence 1137, App
27	45.6	4.1	837	4	US-08-998-416-288	Sequence 288, App



TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
US-09-354-243B-9

Query Match 47.1%; Score 525.2; DB 4; Length 1111;  
Best Local Similarity 73.8%; Pred. No. 8.8e-12;  
Matches 813; Conservative 0; Mismatches 253; Indels 36; Gaps 10;

QY 15 tctgcaatggccgcctcgcagaaatctgtgagctctttctcttatgggagccctggccacc 74  
DB 44 tcgggtgagctgtcctgcagaaatctgtgagctctttctcttatgggagccctggccacc 103

QY 75 agctgctctctctcttcttggccctctgtgtacagggagagcagctgcgccatcagctcc 134  
DB 104 agctgctgctctctcttcttggccctctgtgtacagggagagcagctgcgccatcagctcc 163

QY 135 caactgagcttgacaagcttccaaacttccagagcccttatataccaaacgcaccttctatg 194  
DB 164 cgggtgcaagcttgaggtgtccaaacttccagagcccttatataccaaacgcaccttctatg 223

QY 195 ctgggttaagagagctagctgtgctgatacaacacagagcttctctcattggggagaaa 254  
DB 224 ctgggttaagagagctagctgtgctgatacaacacagagcttctctcattggggagaaa 283

QY 255 ctgttccacggagctcagtagttagtgagcgtctatctgataagcagctgctgaacttc 314  
DB 284 ctgttccacggagctcagtagttagtgagcgtctatctgataagcagctgctgaacttc 343

QY 315 acccttgaagaagtgctgttccctcaatgtataggtttccagccttatatgcagaggtg 374  
DB 344 acccttgaagaagtgctgttccctcaatgtataggtttccagccttatatgcagaggtg 403

QY 375 gtgccccttctgctgagctcagcaacagcttaagcagctgcataatgaaagtgaatc 434  
DB 404 gtgccccttctgagcaacagcttaagcagctgcataatgaaagtgaatc 463

QY 435 ctgcatactcagagaaatgtgcagagctgagggacacagctgagaaagcttgagagagt 494  
DB 464 cagaacatccagagaaatgtgcagagctgagggagcagctgagaaagcttgagagagc 523

QY 495 ggagagatacaagcaatgtgagaaatgtgagagctgtgttattatctctgagaaatgctgc 554  
DB 524 ggagagatacaagcaatgtgagaaatgtgagagctgtgttattatctctgagaaatgctgc 583

QY 555 atttgaccagagcaagctgaaatgaataactaaaccccttccctgctgagaaataac 614  
DB 584 gctgagagagagagctgaaatgaataactaaaccccttccctgctgagaaataac 643

QY 615 aattagatgcccaagcagatttttttaacaaaggaagagagtgagaaagcctccat 674  
DB 644 aataagatccctgaatggacttttta---ctaaaggaagtgagagctaaagctccac 699

QY 675 catgatgggtgagattccaaataaacccctgctgttagttacaaaggaacccaatgcaact 734  
DB 700 catcatgagaagatttcacatgaacctggtcagttgaaagagaaaatagtgctcaa--g 757

QY 735 ttgtttataagcagagagtgagcttcttaagcatagatattttattgataacatttc 794  
DB 758 ttgtttataagcagagagtgagcttcttaagcatagatattttattgataacatttc 816

QY 795 tgaactgggttctatcacagaaacaaatttttttaataatgtctcttttccat 854  
DB 817 tgtcattgat---aatgcaacagaaaaagtgatgacttttaaaaaattgttt-----g 865

QY 855 aaaaaagattacttccatctcttctttaggggaaaaaaccccttaataagctgtctgttcca 914  
DB 866 aaaggaggttactctctctctctta---gaagaaaaagcctatgtaacttca--tttoca 920

QY 915 taactcagactttatttataaaatgattttatttatttataaagcgtcattttattta 974  
DB 921 taaccaataactttattatgtaagttttatttatttataagata-----cattttattta 974

QY 975 tatcatttttaataatgagatttattatagaaacatcattcogatatattgtctacttgagt 1034  
DB 975 tgtcaggttttaataatgagatttattatagaaaaattatctgattgtgtatttgagta 1034

QY 1035 taaggctaatattgatatattatgacaanaattatagagagctataaactgtttatttgacct 1094  
DB 1035 taagcaaaata---atatttatgataaataactatag----aaacaagatatcttagcct 1087

QY 1095 caataaacacttggatatccta 1116  
DB 1088 taataaacacatgaataatcata 1109

RESULT 4  
US-09-178-973B-7  
; Sequence 7, Application US/09178973B  
; Patent No. 6274710  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Renauld, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
; TITLE OF INVENTION: (TIPS)  
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
; FILE REFERENCE: LUD 5543  
; CURRENT APPLICATION NUMBER: US/09/178,973B  
; CURRENT FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 7  
; LENGTH: 1119  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-178-973B-7

Query Match 47.0%; Score 524.4; DB 4; Length 1119;  
Best Local Similarity 74.3%; Pred. No. 1.4e-121;  
Matches 819; Conservative 0; Mismatches 246; Indels 37; Gaps 11;

QY 15 tctgcaatggccgcctcgcagaaatctgtgagctctttctcttatgggagccctggccacc 74  
DB 46 tctctgaggtgctcctgcagaaatctgtgagctctttctcttatgggagccctggccacc 105

QY 75 agctgctctctctcttcttggccctctgtgtacagggagagcagctgcgccatcagctcc 134  
DB 106 agctgctctctctcttcttggccctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 165

QY 135 caactgagcttgacaagtccaaacttccagcagcccttatataccaaacgcaccttctatg 194  
DB 166 cgggtgcaagcttgaggtgtccaaacttccagcagcccttatataccaaacgcaccttctatg 225

QY 195 ctgggttaagagagctagctgtgctgatacaacacagagcttctctcatttggggagaaa 254  
DB 226 ctgggttaagagagctagctgtgctgatacaacacagagcttctctcatttggggagaaa 285

QY 255 ctgttccacggagctcagtagttagtgagcgtctctatctgataagcagctgtgctgaacttc 314  
DB 286 ctgttccacggagctcagtagttagtgagcgtctctatctgataagcagctgtgctgaacttc 345

QY 315 acccttgaagaagtgctgttccctcaatctgattaggttccagccttatatgcagagaggtg 374  
DB 346 acccttgaagaagtgctgttccctcaatctgattaggttccagccttatatgcagagaggtg 405

QY 375 gtgccccttctgagcaggtctcagcaacagcttaagcacatgttcataattgaaagtgatgac 434  
DB 406 gtaccttctgacaaactcagcaatcagctcagctcctgtcactcctgtcacatcagctgacac 465

QY 435 ctgcatactcagagaaatgtgcaaaagctgaaagcacacagctgaaaaagcttggagagagt 494  
DB 466 cagaacatccagagaaatgtgcaaaagctgaaagcacacagctgaaaaagcttggagagagt 525

QY 495 ggagagatcaagcaattggagaactggattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 554



Db 1089 taataaacacatggatcatcata 1110

## RESULT

6

US-09-178-973B-17

: Sequence 17, Application US/09178973B

: Patent No. 6274710

: GENERAL INFORMATION:

: APPLICANT: Dumoutier, Laure

: APPLICANT: Louhed, Jamila

: APPLICANT: Renauld, Jean-Christophe

: TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac

: TITLE OF INVENTION: (TIFs)

: TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof

: FILE REFERENCE: LUD 5543

: CURRENT APPLICATION NUMBER: US/09/178,973B

: PRIOR FILING DATE: 1998-10-26

: NUMBER OF SEQ ID NOS: 17

: SEQ ID NO 17

: LENGTH: 5935

: TYPE: DNA

: ORGANISM: Mus musculus

US-09-178-973B-17

Query Match 19.4%; Score 216.4; DB 4; Length 5935;

Best Local Similarity 69.9%; Pred. No. 1.3e-44;

Matches 446; Conservative 0; Mismatches 156; Indels 36; Gaps 10;

QY 479 aaagcttgagagagtgagagatcaaaatggagaaactggagaaactggattgtgtttatgtc 538

Db 5221 atagcttgagagagcgagagatcaaaagcgtcggaactggagaaactgtgtttatgtc 5280

QY 539 tctgagaaatgctgcatttgaccagagcaaaagctgaaataaataaaccctctt 598

Db 5281 tctgagaaatgctgcatttgaccagagcaaaagctgaaataaataaaccctctt 5340

QY 599 cctgctgagaaatacaaatagatgcccaaaagcgtatttttttaaccaaaaggaagatg 658

Db 5341 gctcttaaaagaacaataaagatccctgaatggactttttt----actaaaggaaagt 5396

QY 659 ggaagccaactccatcatgatgggtggtattccaaatgaacccctgcgttagttacaag 718

Db 5397 agaagtaacgtccacatcattagaaagatttcacatgaacccctgcgttagttacaag 5456

QY 719 gaaaccaatgccactttgtttataagacagaaagtagaacttctaaagcatagatttt 778

Db 5457 aaatagtgctcaa--gtgtccatgagaccag--aggtagacttgataccacaagattc 5513

QY 779 attgataacatttcattgtaactggtgttctatcacagaaaaaacttttttttaaat 838

Db 5514 attgacaatttttttgcattgat----aatgcaacgaaaaagtatgtacttttaaa 5569

QY 839 aattgctttttccataaaaagattactttccattccctttgggggaaaaaacccctaaa 898

Db 5570 aattgcttttttttgcattgat----gaaaggaggttaccctctctctccta---gaaagaaagcctatg 5619

QY 899 tagctctatgtttcccaataacatcagctatttatttataaagtattattattattataa 958

Db 5620 taacttcaa--ttccataaccaactttatataatgtaagttattattattataagata- 5676

QY 959 gactgattttattattattattattattattattattattattattattattattattatt 1018

Db 5677 -----cattttattattattattattattattattattattattattattattattatt 5731

QY 1019 tattgtactgttggtgaaggcttaattattattattattattattattattattattatt 1078

Db 5732 tgttgatttttgatataaagaacata---atattattattattattattattattattatt 5784

QY 1079 catgtttatttgacctcaataaacaacttggatattccta 1116

Db 5785 aagatatcttaggtctttaaaacacatgaatcatcata 5822

## RESULT

7

US-09-354-243B-29

: Sequence 29, Application US/09354243B

: Patent No. 6359117

: GENERAL INFORMATION:

: APPLICANT: Dumoutier, Laure

: APPLICANT: Renauld, Jean-Christophe

: APPLICANT: Louhed, Jamila

: TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible

: TITLE OF INVENTION: (TIFs)

: TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof

: FILE REFERENCE: LUD 5543.1

: CURRENT APPLICATION NUMBER: US/09/354,243B

: PRIOR FILING DATE: 1999-07-16

: CURRENT APPLICATION NUMBER: US/09/178,973

: PRIOR FILING DATE: 1998-10-26

: NUMBER OF SEQ ID NOS: 29

: SEQ ID NO 29

: LENGTH: 5935

: TYPE: DNA

: ORGANISM: Homo sapiens

: FEATURE:

US-09-354-243B-29

Query Match 19.4%; Score 216.4; DB 4; Length 5935;

Best Local Similarity 69.9%; Pred. No. 1.3e-44;

Matches 446; Conservative 0; Mismatches 156; Indels 36; Gaps 10;

QY 479 aaagcttgagagagtgagagatcaaaatggagaaactggagaaactggattgtgtttatgtc 538

Db 5221 atagcttgagagagcgagagatcaaaagcgtcggaactggagaaactgtgtttatgtc 5280

QY 539 tctgagaaatgctgcatttgaccagagcaaaagctgaaataaataaaccctctt 598

Db 5281 tctgagaaatgctgcatttgaccagagcaaaagctgaaataaataaaccctctt 5340

QY 599 cctgctgagaaatacaaatagatgcccaaaagcgtatttttttaaccaaaaggaagatg 658

Db 5341 gctcttaaaagaacaataaagatccctgaatggactttttt----actaaaggaaagt 5396

QY 659 ggaagccaactccatcatgatgggtggtattccaaatgaacccctgcgttagttacaag 718

Db 5397 agaagtaacgtccacatcattagaaagatttcacatgaacccctgcgttagttacaag 5456

QY 719 gaaaccaatgccactttgtttataagacagaaagtagaacttctaaagcatagatttt 778

Db 5457 aaatagtgctcaa--gtgtccatgagaccag--aggtagacttgataccacaagattc 5513

QY 779 attgataacatttcattgtaactggtgttctatcacagaaaaaacttttttttaaat 838

Db 5514 attgacaatttttttgcattgat----aatgcaacgaaaaagtatgtacttttaaa 5569

QY 839 aattgctttttccataaaaagattactttccattccctttgggggaaaaaacccctaaa 898

Db 5570 aattgcttttttttgcattgat----gaaaggaggttaccctctctccta---gaaagaaagcctatg 5619

QY 899 tagctctatgtttcccaataacatcagctatttatttataaagtattattattattataa 958

Db 5620 taacttcaa--ttccataaccaactttatataatgtaagttattattattataagata- 5676

QY 959 gactgattttattattattattattattattattattattattattattattattattatt 1018

Db 5677 -----cattttattattattattattattattattattattattattattattattatt 5731

QY 1019 tattgtactgttggtgaaggcttaattattattattattattattattattattattatt 1078

Db 5732 tgttgatttttgatataaagaacata---atattattattattattattattattattatt 5784

QY 1079 catgtttatttgacctcaataaacaacttggatattccta 1116

Db 5785 aagatatcttaggtctttaaaacacatgaatcatcata 1116

Mon Sep 23 09:43:55 2002

Db 5785 aagatatcttaggctttaataaacaacatgaatatcata 5822

RESULT 8  
US-09-178-973B-8  
; Sequence 8, Application US/09178973B  
; Patent No. 6274710  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renaud, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac  
; TITLE OF INVENTION: (Tifs)  
; FILE REFERENCE: LUD 5543  
; CURRENT APPLICATION NUMBER: US/09/178,973B  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 17  
; SEQ ID NO 8  
; LENGTH: 7445  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-178-973B-8

Query Match 19.3%; Score 215.6; DB 4; Length 7445;  
Best Local Similarity 70.8%; Pred. No. 2.3e-44;  
Matches 452; Conservative 0; Mismatches 149; Indels 37; Gaps 11;  
Qy 479 aaagcttggagagagtgagagagatcaaaagcaattggagaaactgattgctgtttatgctc 538  
Db 6535 atagcttggagagagtgagagagatcaaaagcgattgggaaactggacctgctgtttatgctc 6594  
Qy 539 tctgagaaatcctgcatcttgacagagcaagctgaaagatgaataactaactaaccoccttt 598  
Db 6595 tctgagaaatcctgctgctgagagagagagtagaagcagaaagaaactgctccttctct 6654  
Qy 599 cctgctgagaaatacaactatgagtgcccaagcgtgaaagatgaataactaactaaccoccttt 658  
Db 6655 gccttctaaagaaacataagatccctgaaaggactttttt---actaaaggaaagt 6710  
Qy 659 ggaagccaaactccatcatgagtggtgagatcccaatgaacccctgcgtttagttacaag 718  
Db 6711 agaagctaaagctccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 6770  
Qy 719 gaaaccaaagctccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 778  
Db 6771 aaaaatagtgctcaa--gtgtccatgagacag-aggtgagcttgataaccacaagattc 6827  
Qy 779 attgataacatttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 838  
Db 6828 attgataacatttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 6883  
Qy 839 aattgtctttccataaaagatttccatcttccatcttccatcttccatcttccatcttccatct 898  
Db 6884 aattgtttt-----gaaaggaggttaacctcttccatcttccatcttccatcttccatct 6933  
Qy 899 tagctctcatgtttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 958  
Db 6934 taacttca--tttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 6990  
Qy 959 gactgcatttttttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1018  
Db 6991 -----catttttttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 7045  
Qy 1019 tattgctacttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1078  
Db 7046 tattgata--tttagtataaggcaata---atatattatgacaataactatgagtgagtgag 7097  
Qy 1079 catgtttatttgacctcaataaacaacttgagtgagtgagtgagtgagtgagtgagtgagtgag 1116  
Db 7098 aagatatcttaggctttaataaacaacatgagtgagtgagtgagtgagtgagtgagtgagtgag 7135

RESULT 9  
US-09-354-243B-8  
; Sequence 8, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renaud, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible  
; TITLE OF INVENTION: (Tifs)  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354,243B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: US09/178,973  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 8  
; LENGTH: 7445  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
US-09-354-243B-8

Query Match 19.3%; Score 215.6; DB 4; Length 7445;  
Best Local Similarity 70.8%; Pred. No. 2.3e-44;  
Matches 452; Conservative 0; Mismatches 149; Indels 37; Gaps 11;  
Qy 479 aaagcttggagagagtgagagagatcaaaagcaattggagaaactgattgctgtttatgctc 538  
Db 6535 atagcttggagagagtgagagagatcaaaagcgattgggaaactggacctgctgtttatgctc 6594  
Qy 539 tctgagaaatcctgcatcttgacagagcaagctgaaagatgaataactaactaaccoccttt 598  
Db 6595 tctgagaaatcctgctgctgagagagtagaagcagaaagaaactgctccttctct 6654  
Qy 599 cctgctgagaaatacaactatgagtgcccaagcgtgaaagatgaataactaactaaccoccttt 658  
Db 6655 gccttctaaagaaacataagatccctgaaaggactttttt---actaaaggaaagt 6710  
Qy 659 ggaagccaaactccatcatgagtggtgagatcccaatgaacccctgcgtttagttacaag 718  
Db 6711 agaagctaaagctccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 6770  
Qy 719 gaaaccaaagctccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 778  
Db 6771 aaaaatagtgctcaa--gtgtccatgagacag-aggtgagcttgataaccacaagattc 6827  
Qy 779 attgataacatttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 838  
Db 6828 attgataacatttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 6883  
Qy 839 aattgtctttccataaaagatttccatcttccatcttccatcttccatcttccatcttccatct 898  
Db 6884 aattgtttt-----gaaaggaggttaacctcttccatcttccatcttccatcttccatct 6933  
Qy 899 tagctctcatgtttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 958  
Db 6934 taacttca--tttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 6990  
Qy 959 gactgcatttttttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1018  
Db 6991 -----catttttttccatcatgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 7045  
Qy 1019 tattgctacttgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 1078  
Db 7046 tattgata--tttagtataaggcaata---atatattatgacaataactatgagtgagtgag 7097  
Qy 1079 catgtttatttgacctcaataaacaacttgagtgagtgagtgagtgagtgagtgagtgagtgag 1116  
Db 7098 aagatatcttaggctttaataaacaacatgagtgagtgagtgagtgagtgagtgagtgagtgag 7135



RESULT 10  
US-09-354-243B-25  
; Sequence 25, Application US/09354243B  
; Patent No. 6359117  
; GENERAL INFORMATION:  
; APPLICANT: Dumoutier, Laure  
; APPLICANT: Louhed, Jamila  
; APPLICANT: Renauld, Jean-Christophe  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Factors  
; TITLE OF INVENTION: (TIFFS)  
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof  
; FILE REFERENCE: LUD 5543.1  
; CURRENT APPLICATION NUMBER: US/09/354.243B  
; CURRENT FILING DATE: 1999-07-16  
; PRIOR FILING DATE: 1998-10-26  
; NUMBER OF SEQ ID NOS: 29  
; SEQ ID NO 25  
; LENGTH: 4797  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-354-243B-25

Query Match 18.5%; Score 207; DB 4; Length 4797;  
Best Local Similarity 100.0%; Pred. No. 2.7e-42;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 tcgagttgaaattgtctgcaatggccgcctgcagaaatctgtgagctcttccattgg 60  
Db 52 tcgagttgaaattgtctgcaatggccgcctgcagaaatctgtgagctcttccattgg 111  
Qy 61 ggaacctgaccacagctgcctctcttggccctctgtgacagggagagcagctg 120  
Db 112 ggaacctgaccacagctgcctctcttggccctctgtgacagggagagcagctg 171  
Qy 121 cgcccatcagctcccaactgcagggctgtgacagtcacaaactccagagccctatatcacc 180  
Db 172 cgcccatcagctcccaactgcagggctgtgacagtcacaaactccagagccctatatcacc 231  
Qy 181 accgacacctcagctgctgctaaggagg 207  
Db 232 accgacacctcagctgctgctaaggagg 258

RESULT 11  
US-09-056-075-1/c  
; Sequence 1, Application US/09056075  
; Patent No. 595368  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Eric A.  
; APPLICANT: Bradshaw, Marite  
; APPLICANT: Rood, Julian  
; TITLE OF INVENTION: Expression System for Clostridium  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; CITY: Madison  
; STATE: WI  
; COUNTRY: US  
; ZIP: 53701-2113  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/056.075

FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Seay, Nicholas J.  
REGISTRATION NUMBER: 27386  
REFERENCE/DOCKET NUMBER: 960296.95238  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 608-251-5000  
TELEFAX: 608-251-9166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6243 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 3770..4013  
OTHER INFORMATION: /note= "RP4 origin of DNA transfer (oriT) from  
OTHER INFORMATION: plasmid RP4"  
US-09-056-075-1

Query Match 5.7%; Score 63.2; DB 2; Length 6243;  
Best Local Similarity 48.9%; Pred. No. 1.9e-06;  
Matches 170; Conservative 0; Mismatches 178; Indels 0; Gaps 0;  
Qy 752 aggtgagcttctcagcagatagatattattgtgataacattctgttaactggtgtcttat 811  
Db 1543 AAGTACCAATAAAACATTAGAACGCTCATATTACGTATATATGAAAAAATGGCTTTTAT 1484  
Qy 812 acacagaaacaactatttttttaataattgtcttttccataaaaaagattacttcc 871  
Db 1483 AAGGCTCATTTTTTATATCTTTCTCAAGATATATATATAATAAAAAATTTTTC 1424  
Qy 872 attccttaggggaaacccctaaatagcttcttccataatcagtcacttatat 931  
Db 1423 AAACCTTAAATAAAAAATATTTTATATTTTATTTTATTTTATTTTATTTTATTTTAT 1364  
Qy 932 ttataaagtattattattattataagactgcattttattattattattattattatt 991  
Db 1363 TTTTATTTTATATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1304  
Qy 992 ggaattattatagaacacattcgcattctgctgagtgtaaggctaatatgata 1051  
Db 1303 ATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1244  
Qy 1052 ttatgacaataattatagagctataacatgtttatttgacctcaata 1099  
Db 1243 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1196

RESULT 12  
US-08-487-826B-13/c  
; Sequence 13, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US

ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,826B  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH121.001CPI  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19124 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-487-826B-13

Query Match 5.1%; Score 56.8; DB 2; Length 19124;  
Best Local Similarity 48.7%; Pred. No. 0.00011;  
Matches 154; Conservative 0; Mismatches 162; Indels 0; Gaps 0;

Qy 779 atgataacatttcattgtaactggtgtctctacacagaaacaaattttttttaaattttttt 838  
Db 15773 ATTTTAAATTTTATTTTATGATATATATTTTATTTTAAATATATTTTCTT 15714  
Qy 839 aattgtcttttccataaaaaagattcttccattctttaggggaaaaaacccctaaa 898  
Db 15713 TTTTGTGTTTATGATATATATTTTATTTTAAATGTTTTTTTCTCTTTT 15654  
Qy 899 tagcttcattgtccataacagtaactttattataaaatgtatttattattataaa 958  
Db 15653 GTTTTATTTTATTAATCAATTTTATTTTATATATAATTTTATTTTAAATTTT 15594  
Qy 959 gactgcattttatttattcattttattataatgattttattatagaacattcga 1018  
Db 15593 GATATCTTTTCAATTTTATTTATCTATCAAAATTTATATTTTATTTTATTTATT 15534  
Qy 1019 tattgtacttgagtgtaaggctaatattgatattttatgacaataattatagagctaaa 1078  
Db 15533 TTTTAAAAAATTTTCTCTTTTATTTTATTTTATTTTATTTTATTTTAAATTTT 15474  
Qy 1079 catgtttatttgacct 1094  
Db 15473 ATATTTCAATTTTCTT 15458

RESULT 13  
US-08-443-639-7/c  
Sequence 7, Application US/08443639  
Patent No. 5981843  
GENERAL INFORMATION:  
APPLICANT: Chappell, Joseph  
APPLICANT: Yin, Shaohui  
APPLICANT: Cornett, Catherine A.G.  
TITLE OF INVENTION: Transcriptional Control Sequences and  
METHODS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee and Winner, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder

STATE: CO  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/443,639  
FILING DATE: 18-MAY-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33878  
REFERENCE/DOCKET NUMBER: 69-94  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4254 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Nicotiana tabacum  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1217..1327, 1455..1718, 1806..2182,  
LOCATION: 2259  
LOCATION: ..2477, 2609..2747, 2903..3148, 3262..3558)  
US-08-443-639-7

Query Match 4.6%; Score 51.8; DB 2; Length 4254;  
Best Local Similarity 49.0%; Pred. No. 0.0011;  
Matches 166; Conservative 0; Mismatches 172; Indels 1; Gaps 1;

Qy 690 ccaatgaacccctgcttagttacaaaggaaacccaatgccactttgtttcataagacca 749  
Db 560 CAAGTGTGTTGCTTTGTTTCTCCACACAAAAAATGAAATTTTAAATATATTAATT 501  
Qy 750 gaaggtagactttctaaagcatagatattttatgataacatttcattgtaactgggtct 809  
Db 500 TCAAGTTAGTTTGAATTTAAATTTCTTTTGGTATCGCTCTAGTTTATATTTGTTTT 441  
Qy 810 atacacagaaacaaatttttttaataattgtctttttccataaaaaagattacctt 869  
Db 440 GATGACATAGGTATTAATGTATCTTTTACGCCCTTTTATTTTCAAGAAAAAAGTTGCAT 381  
Qy 870 ccattccttaggggaaaaaacccctaaatagcttcattgtttccataaactcagtaacct 929  
Db 380 GGTGCGGCGCATGTACACACACACACATACATTTATATATATATATATATATATAT 321  
Qy 930 atttataaattgtattattattattataagaactgcattttattttattttattttat 989  
Db 320 ATATATATATATATA-TATATATATATATATATATATATATATATATATATATAT 262  
Qy 990 atggatttattatagaacatcatctcgatattgctact 1028  
Db 261 CTTTTCACAATTTTAGTGTATTTTCTGATTATATATAAT 223

RESULT 14  
US-08-446-855A-1  
Sequence 1, Application US/08446855A  
Patent No. 5849573  
GENERAL INFORMATION:  
APPLICANT: Stewart, Thomas S

APPLICANT: Flores, Maria V  
APPLICANT: O'Sullivan, William J  
TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl  
TITLE OF INVENTION: phosphate synthetase II  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon & Vanderhye PC  
STREET: 1100 No. 5849573th Glebe Road, 8th Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446.855A  
FILING DATE: 06-Jul-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mitchard, Leonard C  
REGISTRATION NUMBER: 29.009  
REFERENCE/DOCKET NUMBER: 47-80  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4100  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8920 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic  
US-08-446-855A-1

Query Match 4.6%; Score 51.8; DB 2; Length 8920;  
Best Local Similarity 53.1%; Pred. No. 0.0015;  
Matches 110; Conservative 0; Mismatches 97; Indels 0; Gaps 0;  
QY 862 attactttccattccttttaggggaaaaaaccccccataagctgctcattggtccataatcag 921  
DB 8452 ATTAACCTTCATCCATAACAACATCAAGTATAAATATATAAATAGTAATATAATA 8511  
QY 922 tactttattataaagtattattattattataaagactgctcattggtccataatcatt 981  
DB 8512 TATAATACGTTAA 8571  
QY 982 ttattaatggattattattattagaaacacatcctgattgctacttgagtgtaaggct 1041  
DB 8572 ATATTAATAAATGTTTTTATTAATATGATCATTAATTATATATGATTATTTTATA 8631  
QY 1042 aatattgatattatgacaaataattat 1068  
DB 8632 AATTTTGTATATATACAAATTTTAT 8658

RESULT 15  
US-09-150-741-1  
Sequence 1, Application US/09150741  
Patent No. 6183996  
GENERAL INFORMATION:  
APPLICANT: Stewart et al.  
TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate  
PATENT NO. 6183996  
TITLE OF INVENTION: Synthetase II  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/09/150,741  
CURRENT FILING DATE: 1998-09-10  
EARLIER APPLICATION NUMBER: PL6380  
EARLIER FILING DATE: 1992-12-16

EARLIER APPLICATION NUMBER: AU93/00617  
EARLIER FILING DATE: 1993-12-02  
EARLIER APPLICATION NUMBER: 08/446,855  
EARLIER FILING DATE: 1995-07-06  
NUMBER OF SEQ ID NOS: 15  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 8920  
TYPE: DNA  
ORGANISM: Plasmodium falciparum  
US-09-150-741-1

Query Match 4.6%; Score 51.8; DB 4; Length 8920;  
Best Local Similarity 53.1%; Pred. No. 0.0015;  
Matches 110; Conservative 0; Mismatches 97; Indels 0; Gaps 0;  
QY 862 attactttccattccttttaggggaaaaaaccccccataagctgctcattggtccataatcag 921  
DB 8452 attactttccattccttttaggggaaaaaaccccccataagctgctcattggtccataatcag 8511  
QY 922 tactttattataaagtattattattattataaagactgctcattggtccataatcatt 981  
DB 8512 TATAATACGTTAA 8571  
QY 982 ttattaatggattattattattagaaacacatcctgattgctacttgagtgtaaggct 1041  
DB 8572 atattaataaagtgttttttataaataatgacataattattattattattatttttata 8631  
QY 1042 aatattgatattatgacaaataattat 1068  
DB 8632 aattttgttatataacaaattttat 8658

Search completed: September 23, 2002, 00:45:14  
Job time: 4335 sec

\_\_\_\_\_

10/23/02

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 23:30:13 ; Search time 1735.48 Seconds  
(without alignments)  
8679.214 Million cell updates/sec

Title: US-09-746-375-1  
Perfect score: 1116  
Sequence: 1 tcgagtagaattgtctgca.....ataaacacttgatccta 1116

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	257.8	23.1	700	11 AK005228	Mus muscu
2	90	8.1	256	9 AV030414	AV030414 AV030414
3	82.6	7.4	1101	12 CNS00EVL	AL069706 Drosophil
4	77.2	6.9	1101	12 CNS0039C	AL063921 Drosophil
5	74.6	6.7	1043	12 CNS0145P	AL103735 Drosophil
6	73.2	6.6	1001	12 CNS0155H	AL105023 Drosophil
7	73.2	6.6	1101	12 CNS00BO1	AL057419 Drosophil
8	71.2	6.4	1101	12 CNS00EVL	AL069706 Drosophil
9	71	6.4	944	12 BH163008	BH163008 ENTF38TR
10	70.4	6.3	389	12 AQ104025	AQ104025 HS_3108_B
11	69.2	6.2	1092	12 CNS020K7	AL175696 Tetraodon
12	69	6.2	1225	12 CNS0161D	AL101671 Drosophil
13	68.6	6.1	1101	12 CNS00088	AL063632 Drosophil
14	68.2	6.1	639	12 CNS038CX	AL232458 Tetraodon
15	68.2	6.1	996	12 CNS00FUH	AL071063 Drosophil
16	67.8	6.1	963	9 AL566565	AL566565 AL566565
17	67.8	6.1	1092	12 CNS020K7	AL175696 Tetraodon

18	67.6	6.1	832	12 CNS01LOK	AL100526 Drosophil
19	67.6	6.1	928	12 CNS00DKY	AL071865 Drosophil
20	67	6.0	843	12 CNS00C51	AL059666 Drosophil
21	67	6.0	1101	12 CNS017KE	AL108152 Drosophil
22	66.2	5.9	842	12 CNS00ACO	AL055007 Drosophil
23	66.2	5.9	895	12 CNS06FSV	AL396821 T7 end of
24	65.8	5.9	928	12 CNS00DKY	AL071865 Drosophil
25	65.8	5.9	1101	12 CNS00FYG	AL071206 Drosophil
26	65.8	5.9	1101	12 CNS016LI	AL106896 Drosophil
27	65.6	5.9	895	12 CNS06FSV	AL396821 T7 end of
28	65.2	5.8	905	12 CNS00KHX	AL077798 Drosophil
29	65	5.8	307	12 CNS00A3W	AL054893 Drosophil
30	64.8	5.8	912	12 AZ551630	AZ551630 ENTF37TF
31	64.8	5.8	1200	12 CNS016CO	AL106578 Drosophil
32	64.6	5.8	1167	12 CNS07360	AL427102 clone BA0
33	64.4	5.8	851	12 BH149587	BH149587 ENTFB44TR
34	64.2	5.8	524	12 CNS01U90	AL167541 Tetraodon
35	64.2	5.8	1190	12 CNS02ON7	AL206908 Tetraodon
36	64	5.7	987	12 CNS014PO	AL104456 Drosophil
37	64	5.7	1101	12 CNS00EO7	AL069440 Drosophil
38	64	5.7	1101	12 CNS00FYG	AL071206 Drosophil
39	63.8	5.7	1101	12 CNS00CYH	AL060100 Drosophil
40	63.6	5.7	845	12 AQ745537	AQ745537 HS_2272_A
41	63.6	5.7	1137	12 CNS071NS	AL423150 clone BA0
42	63.4	5.7	874	12 AZ669565	AZ669565 ENTFM69TF
43	63.2	5.7	687	12 BH133925	BH133925 ENTFNS77TR
44	63.2	5.7	879	12 BH164743	BH164743 ENTRG84TF
45	63.2	5.7	897	12 BH151924	BH151924 ENTFP95TF

## ALIGNMENTS

### RESULT 1

AK005228	AK005228	700 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	Mus musculus adult male cerebellum cDNA, RIKEN full-length enriched library, clone:1500012D04:interleukin 10-related T cell-derived inducible factor, full insert sequence.				
DEFINITION					
ACCESSION	AK005228				
VERSION	AK005228.1	GI:12837639			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult male cerebellum cDNA to mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library				
ORGANISM	clone:1500012D04.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (sites)				
TITLE	Carninci, P. and Hayashizaki, Y.				
JOURNAL	High-efficiency full-length cDNA cloning				
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)				
PUBMED	99279253				
REFERENCE	10349636				
AUTHORS	2 (sites)				
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)				
PUBMED	20499374				
REFERENCE	11042159				
AUTHORS	3 (sites)				
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Todawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
JOURNAL	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer				
	Genome Res. 10 (11), 1757-1771 (2000)				

MEDLINE	20530913	QY	477	aaaaagcttggagagagtgagagagatcaaaagcaatggagaaactgattgctgtttatg	536
PUBMED	11076861	Db	97	AAAAAGCTTTGGAGAGCGGAGAGATCAAAGCGATCGGGAGACTGGACCTGCTGTTTATG	156
REFERENCE	4 (sites)	QY	537	tctctgagaaatcctgcatttgcacagagcaaaagctgaaatgaataaactaacccctt	596
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium	Db	157	TCCTCTGAGAANTCCTTGGCGTCTGAGGAGAAAGACTGAGAAAGAAAGAACTGCTCCTTC	216
TITLE	Functional annotation of a full-length mouse cDNA collection	QY	597	ttccctgctagaaataaacaattagatgcccccaagagatttttttaacccaaaggaaga	656
JOURNAL	Nature 409, 685-690 (2001)	Db	217	CTGCTTCTTAAAAAGAACATAAGATCCCTGATGGACTTTT---ACTAAAGGAAG	272
REFERENCE	5 (bases 1 to 700)	QY	657	tgggaagccaaactccatgatggtggtgattccaaatgaacccctgcgttagttacaa	716
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Inotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	Db	273	TGAGAACTAACGTCCACCATCATTAGAAGATTCACATGAACCTCGCTCAGTTGAAAA	332
TITLE	Direct Submission	QY	717	aggaacccaatgccactttgtttataagaccagaagtagactttcttaagcatagatat	776
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resesgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	Db	333	AGAAAATAGTGTCA--GTTGTCCATGAGACCAG-AGGTAGACTTGATACCAACAAAGAT	389
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.	QY	777	ttattgataacatttcattgtaactggtggttctctatcacagaaacaaatattattttaa	836
FEATURES	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGAGAGAGTCCAGAGCTCTTTTTTTTTTTTNN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGCGCGCAATTAATCTGACTTAATTAATTAATCCGCCGCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.	Db	390	TCATTGACAATATTATTATTGTCATGAT---AATGCAACAGAAAAAGTATGACTTTAA	445
SOURCE	Location/Qualifiers	QY	837	ataattgtttttccataaaaaagattactttccattccttaggggaaaaaccccta	896
Source	1. 700	Db	446	AAAATCTTT-----GAAAGGAGGTTACCTCTCATTTCTCTAGAGAAAAG---CCTA	495
misc_feature	/organism="Mus musculus"	QY	897	aatagctctggttccatcaatcagtcactttatttataaattattattattattattatt	956
gene	/strain="C57BL/6J"	Db	496	TGTAACCTCA--TTTCCATACCACTATTATATATATATATATATATATATATATAT	553
misc_feature	/db_xref="MGI:1900981"	QY	957	aagactgcattttatttatcatatttatttaataatgatttatttattagaaacatcttc	1016
misc_feature	/db_xref="taxon:10090"	Db	554	A-----CATTTTATTATGTCAGTTTATTAAATGCTTTATTATTATAGAAAAATTATCT	607
misc_feature	/clone="1500012D04"	QY	1017	gatttgcattgagtgtaagcctaattgattgatttatttatttatttatttatttattt	1076
misc_feature	/sex="male"	Db	608	GATGTTGATATTGAGTATAAGCAATA---ATATTTATGATTAATTAATTAATTAAT	660
misc_feature	/tissue_type="cerebellum"	QY	1077	aactgttttatttgccttcaataaacacttggtgatc 1113	
misc_feature	/clone_lib="RIKEN full-length enriched mouse cDNA library"	Db	661	ACAGATATCTTAGGCTTTAATAAACACATCAATATC 697	
gene	/dev_stage="adult"	RESULT	2		
misc_feature	/gene="Iltif"	AV030414	AV030414	256 bp	linear
misc_feature	/note="data source:MGI, source key:MGI:1355307, evidence:ISS"	LOCUS	AV030414	Mus musculus adult C57BL/6J cerebellum	EST 31-AUG-1999
BASE COUNT	interleukin 10-related T cell-derived inducible factor"	DEFINITION	AV030414	clone 1500012D04, mRNA sequence.	
ORIGIN	255 a 105 c 133 g 207 t	ACCESSION	AV030414		
Query Match	23.1%; Score 257.8; DB 11; Length 700;	VERSION	AV030414.1	G1:4829962	
Best Local Similarity	70.9%; Pred. No. 5.7e-39;	KEYWORDS	EST		
Matches 494; Conservative	0; Mismatches 167; Indels 36; Gaps 10;	SOURCE	house mouse		
QY	417 catattgaagtgcactgcattatccagaggaattgcaaaagctgaagacacagt 476	ORGANISM	Mus musculus		
Db	37 CACATCAGCGGTGACGACCAACATCCAGAGAAATGTCAGAAAGCTGAGAGACAGTG 96	REFERENCE	1 (bases 1 to 256)		
		AUTHORS	Akahira, S., Akiyama, J., Fukuda, S., Fukunishi, Y., Funayama, T., Hara, A., Hayatsu, N., Hori, F., Ishikawa, T., Itoh, M., Izawa, M., Kawai, J., Kikuchi, N., Kojima, Y., Matsuyama, T., Nitsuma, H., Oda, H., Owa, C., Sato, K., Shibata, Y., Shigemoto, Y., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tateno, M., Tomaru, Y., Tomimaga, N., Watanabe, S., Yagame, M., Yamamura, T., Yokota, T., Yoshino, M., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.		
		TITLE	RIKEN Mouse ESTs		
		JOURNAL	Unpublished (1999)		
		COMMENT	Contact: Chie Owa		
			Genome Science Laboratory		
			RIKEN		
			3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan		

Tel: 81-298-36-9145  
 Fax: 81-298-36-9098  
 Email: genome-res@rtc.riken.go.jp  
 Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))  
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))  
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES  
 source  
 Location/Qualifiers  
 1..256  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="1500012D04"  
 /clone\_lib="Mus musculus adult C57BL/6J cerebellum"  
 /sex="male"  
 /tissue\_type="cerebellum"  
 /dev\_stage="adult"  
 BASE COUNT 88 a 29 c 41 g 98 t  
 ORIGIN

Query Match 8.1%; Score 90; DB 9; Length 256;  
 Best Local Similarity 72.1%; Pred. No. 3e-07;  
 Matches 147; Conservative 0; Mismatches 50; Indels 7; Gaps 2;  
 QY 910 tcccaatcagctactttattataatgattattattattattattataagactgatttt 969  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 60 TTCTCTCCATACCGATACCTTTATATATGTCAGTTTATTTATATAAGTATACATTTT 119  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 970 atttatcatctttattataatgattttattataagactgattttattattattattatt 1029  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 120 ATTATGTCAGTTTATTAATATGATTTTATTTATAGAAAATATCTGATGATTTT 179  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1030 gaggtaaggcctaattgattattatgacaataattatagactataacattgtttattt 1089  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 180 GAGTATAAGCAAAATA--ATATTTATGATAATACTATAG----AAACAAGATATCTTA 232  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1090 gacccaataaacacttgatc 1113  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 233 GCGTTTAAACACATGATATC 256  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3  
 CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC: BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION AL069706  
 VERSION AL069706.1 GI:4949849  
 KEYWORDS GSS:  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 1101)  
 AUTHORS Genoscope.  
 DIRECT SUBMISSION  
 TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr))  
 JOURNAL - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Hammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
 Location/Qualifiers  
 1..1101  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone\_lib="RPCI-98"  
 /clone="BACR29B23"  
 /note="end : 77"  
 BASE COUNT 419 a 91 c 60 g 299 t 232 others  
 ORIGIN

Query Match 7.4%; Score 82.6; DB 12; Length 1101;  
 Best Local Similarity 35.7%; Pred. No. 7.7e-06;  
 Matches 139; Conservative 78; Mismatches 172; Indels 0; Gaps 0;  
 QY 713 acaaaggaaacccaatgccacttttggttataagaccagaggtagactttctaagcatag 772  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 698 AAAAAAAWATAWAAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATA 757  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 773 atatttattgaaacatttccattgtaactggtctctatcacagacaaacattttttt 832  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 758 WATATATATWTTTAAWATWAAWTTATWATAWATAWATAWATAWATAWATAWATAWATA 817  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 833 ttaataattgtcttttccataaaaaagattacttccattcccttaggggaaaaaac 892  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 818 TWAAAWAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATA 877  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 893 cctaaatagcttcagtttccataatcagctactttattattataaattgtatttattatta 952  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 878 AATAWATWTTWTTTWWAAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATAW 937  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 953 ttataagactgctatttattattattattattattattattattattattattattattatt 1012  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 938 WTWATATWTTTATTAATWTTATWTTATWTTATWTTATWTTATWTTATWTTATWTTATWTT 997  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1013 attcgatttgcacttgagtgtaagcctaattatgattattatgacaataattatagag 1072  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 998 AWTAWATATATTTATTAWTTATTTTAAWAAWATAWATAWATAWATAWATAWATAWATAW 1057  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1073 ctataacatgtttatttgacctcaataaa 1101  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1058 WAATTATWTTATATWTTAAWATAWATAWATAWATAWATAWATAWATAWATAWATAWATA 1086  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4  
 CNS0039G 1101 bp DNA linear GSS 03-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION AL063921  
 VERSION AL063921.1 GI:4941778  
 KEYWORDS GSS:  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 1101)  
 AUTHORS Genoscope.  
 DIRECT SUBMISSION  
 TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr))  
 JOURNAL - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).





RESULT	6
CNS0155H/c	
LOCUS	1001 bp DNA linear GSS 26-JUL-1999
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BRACN13C23 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL105023
VERSION	ALI05023.1 GI:5617037
KEYWORDS	GSS.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 1001)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr. - web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.
FEATURES	Location/Qualifiers
source	1..1001
	/organism="Drosophila melanogaster"
	/plasmid="pBelOBAC11"
	/db_xref="taxon:7227"
	/clone_lib="DrosBAC"
	/clone="BACN13C23"
	/note="end : SP6"
BASE COUNT	266 a 219 c 134 g 150 t 232 others
ORIGIN	
Query Match	6.6%; Score 73.2; DB 12; Length 1001;
Best Local Similarity	40.9%; Pred. No. 0.00046;
Matches 108; Conservative	50; Mismatches 106; Indels 0; Gaps 0;
QY 826	ttattttaaataatgtctttcccaaaaaagattactttccattccttccttagggga 885
I :	::::
Db 982	TTTWTWTTWATTTTTATTATTATTATTATAAATAAATAAATAAATAAATAAATAA 923
QY 886	aaacacccttaaatagcttcagtgttccacaataacgactcatcttatataaatgtattt 945
I :	::::
Db 922	AWAHHTTWTAAWWAAATTTTWAATTTTATTTTWTTTTTTTTTTTTATATATWATTW 863
QY 946	attattattaagaactgcattttatttatatcatctttttaataatgatgtatttatttag 1005
I :	::::
Db 862	WTTWTWTTTAAAAAATAAATTTTITTTTWTAAWTTTWTAAAATTTTWTATTTTWTWT 803
QY 1006	aacatcatcgatctgacttgctgagtgtaaggctaataatgatattatgatgacaaaat 1065
I :	
Db 802	AAATATTTTTTAAWTTTTTTTTTATTATTAATAAATAAATAAATAAATAAATAA 743
QY 1066	tatagagctataacatgtttattt 1089
I :	
Db 742	AATTTTTTTHVYYTWYTATT 719
RESULT	7
CNS00BO1	
LOCUS	1101 bp DNA linear GSS 04-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence W7 end of BAC # BACR24D09 of RCI198 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	ALJ057419
VERSION	ALJ057419.1 GI:4937885

LOCUS  
 DEFINITION  
 Drosophila melanogaster genome survey sequence T7 end of BAC:  
 BACR29B23 of RPI-98 library from Drosophila melanogaster (fruit  
 fly), genomic survey sequence.  
 AL069706  
 VERSION  
 AL069706.1 GI:4949849  
 GSS.  
 SOURCE  
 Drosophila melanogaster  
 fruit fly.  
 ORGANISM  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE  
 1 (bases 1 to 1101)  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
 - web : www.genoscope.cns.fr)  
 COMMENT  
 Determination of this BAC-end sequence was carried out as part of a  
 collaboration with the Berkeley Drosophila Genome Project (BDGP).  
 The BDGP is constructing a physical map of the Drosophila  
 melanogaster genome using these BACs. For further information  
 please see <http://www.fruitfly.org> The BDGP Drosophila  
 melanogaster BAC library was prepared by Kazutoyo Osogawa and  
 Aaron Mamoser in Pletier de Jong's laboratory in the Department of  
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,  
 NY. The library is named RPI-98 and was constructed by partial  
 EcoRI digestion of Drosophila DNA provided by the BDGP from the  
 isogenic strain Y2: cn bw sp, the same strain used for the BDGP's  
 P1 and EST libraries. A more detailed description of the library  
 and how to order individual BAC clones, the entire library, or  
 filters for hybridization from the BACPAC Resource Center can be  
 found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES  
 source  
 Location/Qualifiers  
 1..1101  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone\_lib="RPI-98"  
 /clone="BACR29B23"  
 /note="end : T7"  
 BASE COUNT 419 a 91 c 60 g 299 t 232 others  
 ORIGIN  
 Query Match 6.4%; Score 71.2; DB 12; Length 1101;  
 Best Local Similarity 35.8%; Pred. No. 0.0011;  
 Matches 118; Conservative 67; Mismatches 145; Indels 0; Gaps 0;  
 Qy 760 ttcttaagcatagattatttattgatacaatttcattgtaactggtgtctctacacagaa 819  
 Db 1085 TTTTWTATTTTAAATATATATATATTTTWTATATATATATATATATATATATATAT 1026  
 Qy 820 acaattatttttttaataattgctctttccataaaagattacttccattcccttt 879  
 Db 1025 AAAATATATATTTTAAATATATATATATATATATATATATATATATATATATAT 966  
 Qy 880 aggggaaaaaaccccaaatagcttcattgttcccaatacagtaactttattttataaat 939  
 Db 965 AATWATATATTTTAAATAAATATATATATATATATATATATATATATATATATAT 906  
 Qy 940 gkattattattattataagacgcattttatttatacaattttattataatgagattat 999  
 Db 905 TTATATTTTAAAT 846  
 Qy 1000 ttataaacatcatctgattgtctactgtgagtgtaaggctaatattgatattatgatc 1059  
 Db 845 ATTWTATATATTTTAAATATATATATATATATATATATATATATATATATATAT 786  
 Qy 1060 aataattatagagcatataacattgtttattt 1089  
 Db 785 TATATTTTAAAT 756

RESULT 9  
 BH163008/c  
 LOCUS  
 DEFINITION  
 Entamoeba histolytica Sheared DNA Entamoeba histolytica  
 genomic, DNA sequence.  
 BH163008  
 ACCESSION  
 BH163008.1 GI:15736446  
 GSS.  
 VERSION  
 BH163008.1  
 KEYWORDS  
 SOURCE  
 Entamoeba histolytica.  
 ORGANISM  
 Entamoeba histolytica  
 Eukaryota; Entamoebidae; Entamoeba.  
 REFERENCE  
 1 (bases 1 to 944)  
 AUTHORS  
 Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.  
 TITLE  
 Determination of clone end sequences from Entamoeba histolytica  
 HMI:IMSS sheared DNA library (2001)  
 JOURNAL  
 Unpublished (2001)  
 COMMENT  
 Contact: Brendan J Loftus  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0208  
 Fax: 301 838 3543  
 Email: bjl@loftus@tigr.org  
 Clones are derived from the Entamoeba histolytica HMI:IMSS sheared  
 DNA library  
 Seq primer: M13-Reverse  
 Class: shotgun  
 High quality sequence start: 13  
 High quality sequence stop: 559.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..944  
 /organism="Entamoeba histolytica"  
 /strain="HMI:IMSS"  
 /db\_xref="taxon:5759"  
 /clone\_lib="Entamoeba histolytica Sheared DNA"  
 /note="Vector: PHOS1; Site 1: Bst I; Constructed at The  
 Institute for Genomic Research (TIGR), Rockville, MD.  
 Genomic DNA isolated from broth cultures of E. histolytica  
 using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993). Entamoeba histolytica: a  
 method for isolate identification. Exp. Parasitol.  
 77:450.). The DNA was mechanically sheared to give a  
 tight size distribution (~2 kb). The v + i method used for  
 the library construction is described in detail in Smith,  
 H.O. and Venter, J.C. (Making small insert libraries for  
 whole genome shotgun sequencing projects. In Genome  
 Sequencing: A Practical Approach, eds. M. Vaudin and B.  
 Barrell, Oxford University Press, 1999)."

BASE COUNT 430 a 67 c 197 g 250 t  
 ORIGIN  
 Query Match 6.4%; Score 71; DB 12; Length 944;  
 Best Local Similarity 50.4%; Pred. No. 0.0012;  
 Matches 173; Conservative 0; Mismatches 170; Indels 0; Gaps 0;  
 Qy 773 atatttattgatacaatttcattgtaactggtgtctctacacagaaacaaatttttt 832  
 Db 594 ATACTTATATATCTCTATATACCTTTATATATATATATATATATATATATATATAT 535  
 Qy 833 ttaataatctgctctttccataaaagattacttccattcccttttaggggaaaaaac 892  
 Db 534 TTAATAACTCTAT 475  
 Qy 893 cctaaatagcttcattgttcccaataatcagtaactttattataatgattattattat 952  
 Db 474 TTTATATGTTTAT 415  
 Qy 953 ttataagactgcatttttatttatacatattttattataatgagattttattatagaacac 1012  
 Db 414 TATGTTTATATGTTTATATATATATATATATATATATATATATATATATATAT 355  
 Qy 1013 attcgatattgctacttgagtgtagggcctaattatgatattttatgacaataattatagag 1072

	CNS0161D	DNA	linear	GSS 26-JUL-1999
	LOCUS	1225 bp		
	DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of HAC		
sequence.				
ACCESSION	ALL75696			
VERSION	ALL75696.1			GI:7813753

BACN15C18 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION AL106171  
 VERSION AL106171.1 GI:5620504  
 KEYWORDS GSS.  
 SOURCE fruit fly.  
 ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 1225)  
 AUTHORS Direct Submission  
 TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :  
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)  
 COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES Location/Qualifiers  
 source 1..1225

BASE COUNT 266 a 128 c 38 g 368 t 425 others  
 ORIGIN

Query Match 6.2%; Score 69; DB 12; Length 1225;  
 Best Local Similarity 31.6%; Pred. No. 0.0029;  
 Matches 135; Conservative 93; Mismatches 199; Indels 0; Gaps 0;  
 Oy 605 tagaaatacaaatgagtcgcccaagcatttttttaaccacaaagaaagatgggaagc 664  
 Db 774 KAAAMNAAHHWTTTYYAAAYNNAANAATTTTAAATCKCAAMAAWAAWAAWAAANAC 833  
 Oy 665 caaacctcatcatgatgggtgattcccaagaaacccctgcgttagttacaaagaaacc 724  
 Db 834 CMCVCCKKKKKKKKKKKKKDDAKRARADRRKAAWAAWAAWAAWAAWAAWAAWAAW 893  
 Oy 725 aatgcacattttgtttataagaccagaaggtagacctttcttaagcatagattttatgat 784  
 Db 894 AAAAMWATTTTYYAYAWAAAYAWATAWAAWTTTAAWAAWAAWAAWAAWAAWAAWTTTY 953  
 Oy 785 acatttcattgtaactgggtgtctctacacagaaacacatttttttaaaataattgt 844  
 Db 954 AAAAAMWTTWAAATPAWTTATTTTATTTTAAWTTTATWATAATTTATWTTTATTTT 1013  
 Oy 845 ctttttccataaaaagagacttctccattcttaggggaaacccccaataagctt 904  
 Db 1014 TTTTAAWAAWAAWAAWAAWAAWAAWAAWTTTAAWTTTATATATATATATATATATAT 1073  
 Oy 905 catgtttccataaatcagctactttattataaaatgtattttattattataagactgc 964  
 Db 1074 ATTTAAWAAWAAWTTTWT 1133  
 Oy 965 attttattatactattttataaataatggattttattataagaaacattcogactatgc 1024  
 Db 1134 ATAAAAAATTTTWTWAAATAAWATATATATATATATATATATATATATATATATATAT 1193  
 Oy 1025 tacttga 1031  
 Db 1194 WTWTWAA 1200

RESULT 13  
 CNS000B8  
 LOCUS  
 DEFINITION

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

FEATURES source  
 Location/Qualifiers

BASE COUNT 228 a 114 c 110 g 512 t 137 others  
 ORIGIN

Query Match 6.1%; Score 68.6; DB 12; Length 1101;  
 Best Local Similarity 39.8%; Pred. No. 0.0034;  
 Matches 134; Conservative 47; Mismatches 156; Indels 0; Gaps 0;

Oy 755 tagactttctaaagcagatagattttattgatacaatttcattgtaactgggtgtctatata 814  
 Db 734 TWTATWTTWTTWAAWAWTTTWTWTTTATTTTAAATTTWATTTTTCGTAATTTTAAATTTTADT 793  
 Oy 815 cagaaacacatttttttaaaataatgtctttttccataaaaaaagatttcttccatt 874  
 Db 794 TTTTWTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 853  
 Oy 875 cctttagggggaaaaacccctaaatagcttcattgtctccataactcagctacttattata 934  
 Db 854 TTTTAAWAAWAAWAAWAAWAAWAAWAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAA 913  
 Oy 935 taaatgtatttattattatacaagactgcattttattattatcatctttattataatgga 994  
 Db 914 TWTWTTWATWTTWAAWTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 973  
 Oy 995 tttattttatagaacacattcogactgctacttgagtgtaaggcctaataatctgattt 1054  
 Db 974 TTTAAWTTTATTTAAWATATTTATTTTAAWTTTATTTTATTTTATTTTATTTTATTTTATTT 1033  
 Oy 1055 atgacaataattatagagctatacaactgtttatttga 1091  
 Db 1034 AWTWAAWATWATWTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1070

CNS000B8 1101 bp DNA linear GSS 03-JUN-1999  
 Drosophila melanogaster genome survey sequence TET3 end of BAC #  
 BACR01A24 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL063632  
 AL063632.1 GI:4938680  
 GSS.  
 SOURCE fruit fly.

ORGANISM Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 1101)  
 AUTHORS Direct Submission

TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
 - Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila\_bac.hum.

FEATURES Location/Qualifiers  
 source 1..1101

BASE COUNT 228 a 114 c 110 g 512 t 137 others

Query Match 6.1%; Score 68.6; DB 12; Length 1101;  
 Best Local Similarity 39.8%; Pred. No. 0.0034;  
 Matches 134; Conservative 47; Mismatches 156; Indels 0; Gaps 0;

Oy 755 tagactttctaaagcagatagattttattgatacaatttcattgtaactgggtgtctatata 814  
 Db 734 TWTATWTTWTTWAAWAWTTTWTWTTTATTTTAAATTTWATTTTTCGTAATTTTAAATTTTADT 793  
 Oy 815 cagaaacacatttttttaaaataatgtctttttccataaaaaaagatttcttccatt 874  
 Db 794 TTTTWTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 853  
 Oy 875 cctttagggggaaaaacccctaaatagcttcattgtctccataactcagctacttattata 934  
 Db 854 TTTTAAWAAWAAWAAWAAWAAWAAWAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAAWTTTAA 913  
 Oy 935 taaatgtatttattattatacaagactgcattttattattatcatctttattataatgga 994  
 Db 914 TWTWTTWATWTTWAAWTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 973  
 Oy 995 tttattttatagaacacattcogactgctacttgagtgtaaggcctaataatctgattt 1054  
 Db 974 TTTAAWTTTATTTAAWATATTTATTTTAAWTTTATTTTATTTTATTTTATTTTATTTTATTT 1033  
 Oy 1055 atgacaataattatagagctatacaactgtttatttga 1091  
 Db 1034 AWTWAAWATWATWTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1070

```
RESULT 14
CNS038CX
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence T7 end of clone
005C19 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION
AL232458
VERSION
AL232458.1 GI:7891593
KEYWORDS
GSS; genome survey sequence.
SOURCE
Tetraodon nigroviridis.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE
1 (bases 1 to 639)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Boutin,A., Fizames,C., Wincker,P., Brothier,P., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 639)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL
Unpublished
REFERENCE
3 (bases 1 to 639)
Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE
Direct Submission
JOURNAL
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
AUTHORS
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
1..639
/organism="Tetraodon nigroviridis"
/db_xref="taxon:9983"
/clone="005C19"
/clone_lib="G"
/note="Genoscope sequence ID : CORG005AB10XD1-end : T7"
BASE COUNT
210 a 3 c 5 g 381 t 40 others
ORIGIN
1..639
/organism="Tetraodon nigroviridis"
/db_xref="taxon:9983"
/clone="005C19"
/clone_lib="G"
/note="Genoscope sequence ID : CORG005AB10XD1-end : T7"
Query Match 6.1%; Score 68.2; DB 12; Length 639;
Best Local Similarity 46.5%; Pred. No. 0.004;
Matches 148; Conservative 6; Mismatches 164; Indels 0; Gaps 0;
QY 774 tatttatgataacatttcattgtaactgagtggtctctatcacagagaacattttttt 833
Db 43 TATTATTATTATTTTAAACCTTTAATTATTATATAAAAAATAAATAATATAATTAT 102
QY 834 taataaattgtcttttccataaaaagattacttccattcccttaggggaaaccc 893
Db 103 TAAATATTATTTTAAATAAATAAATAATTAATTAATAAATAAATAAATAAATAAATA 162
QY 894 ctaaatagcttcatttccataaatcagtaactttatattataaaatgattattattat 953
Db 163 NNTNNAANNNAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 222
QY 954 tataagactgcattttatttatacatcttttataatagatttttttttttttttttttt 1013
Db 223 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 282
QY 1014 ttcgattatgcactgagtgtaagcgaataattgatattttatgacaataattatagac 1073
Db 283 TTTTATATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATA 342
```

```
QY 1074 tataacatgttttttttga 1091
1:| | | | |
Db 343 AAWAATTTTATATTTTGA 360
1:| | | | |

RESULT 15
CNS00FUH/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR31021 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL071063
VERSION
AL071063.1 GI:4951105
KEYWORDS
GSS.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 996)
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
JOURNAL
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammos in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Location/Qualifiers
1..996
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR31021"
/note="end : TET3"
BASE COUNT
383 a 164 c 81 g 171 t 197 others
ORIGIN
1..996
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR31021"
/note="end : TET3"
Query Match 6.1%; Score 68.2; DB 12; Length 996;
Best Local Similarity 36.7%; Pred. No. 0.0041;
Matches 109; Conservative 61; Mismatches 127; Indels 0; Gaps 0;
QY 804 tcttcatacacagaaacaatttttttaaaatgctcttttccataaaagaat 863
Db 990 TTTTWTATAAAAAATTTTWTATTAATAAATAAATAAATAAATAAATAAATAAATAA 931
QY 864 tacttccattccttaggggaaacccctaaatagcttcattggttccataacagta 923
Db 930 TTTBGGSSSSSSGSSGSSGRRRAAAAAAATAAATAAATAAATAAATAAATAAATAA 871
QY 924 ctttatattataatgtatttattattataaagacgcatctttatttattatcat 983
Db 870 RTTWRATTWTAAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 811
QY 984 attaatgagttattattatagaacattcattgctactgagtglaagagctaa 1043
Db 810 ATWWTMAAANWTTAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 751
QY 1044 tattgatattatgacaataattatagagctataacatgttttattgacctcaataa 1100
```

**D**b 750 TAWTTATTTTTWTAATTTTTWWATAWWWTTTTTATTWTTTTTTTTTTTTTTTTWA 694

Search completed: September 23, 2002, 00:02:37  
Job time: 1944 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 23, 2002, 00:45:19 ; Search time 2215.22 Seconds  
(without alignments)  
10542.536 Million cell updates/sec

Title: US-09-746-375-1  
Perfect score: 1116  
Sequence: 1 tcgagttagaattgtctgca.....ataaacacttgatccta 1116

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 14

Total number of hits satisfying chosen parameters: 151023

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_em.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
------------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1

AX151713

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

AX151713 1116 bp DNA linear PAT 22-JUN-2001

Sequence 14 from Patent WO0140467.

AX151713

AX151713.1 GI:14533647

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1116)

Presnell,S.R., Xu,W., Kindsvogel,W. and Chen,Z.

Human cytokine receptor

Patent: WO 0140467-A 14 07-JUN-2001;

ZymoGenetics, Inc. (US)

Location/Qualifiers

1..1116

/organism="Homo sapiens"

/db\_xref="taxon:9606"

21..560

/note="unnamed protein product"

/codon\_start=1

```
/protein_id="CAC42593.1"
/db_xref="GI:14533648"
/translation="MAALQKSVSFLMGTLATSCILLALLVQGGAAAPISSHCRDLK
SNFQOPYITNRTFMALKEASLADNTDVRLLGKLFHGVMSERCYLMKQVLFNFTLEE
VLFPSDRFPQYMOEVPPLARLSNRLSCHIEGDDLHIQRNVQKLKDPVKKLGESGE
IKAIQELDLFLMSLRNACI"
BASE COUNT 336 a 229 c 222 g 329 t
ORIGIN
Query Match 100.0%; Score 1116; DB 6; Length 1116;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 tcgagttagaattgtctgcaatgcccgcctcgagaaatctgtgagctcttctctatgg 60
Db 1 TCGAGTTAGAAATTGCTGCAATGCCGCCCTCGAGAAATCTGTGAGCTCTTCTCTATGG 60
Qy 61 ggacctggccaccagctgctctctcttcttggccctcttgggtacagggggagcagctg 120
Db 61 GGACCCCTGGCCACGAGCTGCTCTCTCTTGGCCCTCTTGGTACAGGGAGGACAGCTG 120
Qy 121 cgcctacagctccactgagctgacaaagtcacaaacttccagcagccctatatcacca 180
Db 121 CGCCCATCAGCTCCCACTGACAGGCTTGACAAGTCCAACTTCCAGACCCCTATATCACCA 180
Qy 181 accgcaccttcactgctgagcagctgagctgagctgagctgagctgagctgagctg 240
Db 181 ACCGCACCTTCATGCTGAGCTGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 240
Qy 241 tcattggggagaaactgttccacagagctgagctgagctgagctgagctgagctgagc 300
Db 241 TTGCTCTTTTCCATAAAAAAGATTACCTTTCCATTTCCATTTCCATTTCCATTTCCAT 300
```

```
Qy 901 gcttcactgttccatcaatcagtcacttattattattattattattattattattataaga 960
Db 901 GCTTCATGTTCCATATCATAGTACTTATATTATATAATGCTATTTATTATTATTATAAGA 960
Qy 961 ctgcatttattattattattattattattattattattattattattattattattattcgata 1020
Db 961 CTGCATTTTATTATATCATTTTATTATATGATTTATTATTATTATTATTATTATTATTATGATA 1020
Qy 1021 ttgctactgtgagtgtaaggctcaatatttgatattattattattattattattattattataa 1080
Db 1021 TTGCTACTTTCAGTGTAGGCTAATATTGATATTATTATGACAATAATTATTAGAGCTATAACA 1080
Qy 1081 tggttattgacctcaataaacaacttggatataccta 1116
Db 1081 TGTATTATTGACCTCAATAAACACACTTGGATATCCTA 1116
RESULT 2
AX179578
LOCUS 1116 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent WO0146422.
ACCESSION AX179578
VERSION AX179578.1 GI:15132010
KEYWORDS human.
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1116)
AUTHORS Presnell,S.R. and Kindsvogel,W.
TITLE Cytochrome zyto18
JOURNAL Patent: WO 0146422-A 1 28-JUN-2001;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1..1116
/organism="Homo sapiens"
/db_xref="taxon:9606"
21..560
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC49973.1"
/db_xref="GI:15132011"
/translation="MAALQKSVSFLMGTLATSCILLALLVQGGAAAPISSHCRDLK
SNFQOPYITNRTFMALKEASLADNTDVRLLGKLFHGVMSERCYLMKQVLFNFTLEE
VLFPSDRFPQYMOEVPPLARLSNRLSCHIEGDDLHIQRNVQKLKDPVKKLGESGE
IKAIQELDLFLMSLRNACI"
BASE COUNT 336 a 229 c 222 g 329 t
ORIGIN
Query Match 100.0%; Score 1116; DB 6; Length 1116;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 tcgagttagaattgtctgcaatgcccgcctcgagaaatctgtgagctcttctctatgg 60
Db 1 TCGAGTTAGAAATTGCTGCAATGCCGCCCTCGAGAAATCTGTGAGCTCTTCTCTATGG 60
Qy 61 ggacctggccaccagctgctctctcttcttggccctcttgggtacagggggagcagctg 120
Db 61 GGACCCCTGGCCACGAGCTGCTCTCTCTTGGCCCTCTTGGTACAGGGAGGACAGCTG 120
Qy 121 cgcctacagctccactgagctgacaaagtcacaaacttccagcagccctatatcacca 180
Db 121 CGCCCATCAGCTCCCACTGACAGGCTTGACAAGTCCAACTTCCAGACCCCTATATCACCA 180
Qy 181 accgcaccttcactgctgagcagctgagctgagctgagctgagctgagctgagctgagc 240
Db 181 ACCGCACCTTCATGCTGAGCTGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 240
Qy 241 tcattggggagaaactgttccacagagctgagctgagctgagctgagctgagctgagc 300
Db 241 TTGCTCTTTTCCATAAAAAAGATTACCTTTCCATTTCCATTTCCATTTCCATTTCCAT 300
```



Db 241 TCATTGGGAGAACTGTTCCAGGAGTCAGTATGAGTGAGCGCTGCTATCTGATGAAC 300

Qy 301 agtgctgaactcacccttgaagtgctgttccctcaatctgatagttccagcctt 360

Db 301 AGTGCTGAACCTTACCCTTGAAGAGTCTGTTCCCTCAATCTGATAGTTCCAGCCTT 360

Qy 361 atatcagagagtggtgcttccctggcagagctcagcaaacaggctaaagcatgtcata 420

Db 361 ATATCAGAGAGTGTGGCTTCCCTGGCAGGCTCAGCAACAGGCTAAGCATGTCTATA 420

Qy 421 ttgagagtgatgacactgcataccagaggaatgtcgaagagctgaagcacagagaaaa 480

Db 421 TTGAAGTGATGACCTGCATATCCAGAGAAATGTCAAAAGCTGAAGGACAGTGAATA 480

Qy 481 agcttgagagagtgagagatcaagcaatggagagctgatttgcgtttatgtctc 540

Db 481 AGCTTGGAGAGTGGAGAGATCAAGCAATTTGGAGAACTGGATTTGCTGTTATGTCTC 540

Qy 541 tgagaaatgctgcaatttgaccagagaaagctgaaatgaataactaacccccctttcc 600

Db 541 TGAGAAATGCCCTGCATTTGACCAGAGCAAGCTGAAATGAATAACTAACCCCTTTCC 600

Qy 601 ctgctagaataaacaatagatgccccaaagcgatttttttaaccaaagaagatggg 660

Db 601 CTGCTAGAATAACAAATTAGATGCCCAAGCGATTTTTTTTAAACAAAGAAGATGG 660

Qy 661 aagcacaactccatcatgatggtggtgattcccaatgaacccctggttagttacaaaga 720

Db 661 AAGCCAAACTCCATCATGATGGTGGATTCACAAATGAACCCCTGGCTTAGTTACAAAGA 720

Qy 721 aaccaatgcaactttgtttataagacagaaggtagactttcctaagcagatatttat 780

Db 721 AACCAATGCCACTTTGTTTATAAGACCAAGAGTAGACTTTCTAAGCATAGATATTTAT 780

Qy 781 tgataaacttcatgttaactggtgtctctacacagaaacaatttttttaataaa 840

Db 781 TGATAAATTCATGTACTGCTACTGCTCTCTATACAGAAACAATTTATTTTAAATAA 840

Qy 841 ttgctcttttccataaaaaagattactttccattccctttgagggaaaaaacccctaaata 900

Db 841 TTGCTCTTTTCCATAAAAAAGATTACTTTCCATTCCTTTAGGGGAAAAAACCCCTAAATA 900

Qy 901 gcttcatgttccataatacagtaactttattataataatgattattattattataaga 960

Db 901 GCTTCATGTTTCCATAAATCAGTACTTATTTATTAATAATGATTTATTTATTTATAAGA 960

Qy 961 ctgcatatttattatcatcttttattataatgattttattatagaacatcatctcata 1020

Db 961 CTGCAATTTATTATATCATTTTATTATATGATTTATTTATAGAAACATCATTCGATA 1020

Qy 1021 ttgctacttgagtgtaaggctaatttgatatttatgacaataattatagactataaca 1080

Db 1021 TTGCTACTTGTAGTGAAGCTAATATGATTTATGACAAATTTATAGACTATAACA 1080

Qy 1081 tgtttatttgactcaataaaacacttggtatccta 1116

Db 1081 TGTATTATTGACCTCAATAAACACTTGGATATCCTA 1116

RESULT 3

AX092422

LOCUS AX092422

DEFINITION Sequence 153 from Patent WO0116318.

ACCESSION AX092422

VERSION AX092422.1 GI:13444525

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1152)

AUTHORS Eaton,D.L., Filvaroff,E., Gerritsen,M.E., Goddard,A., Godowski,P.J., Grimaldi,C.J., Gurney,A.L., Watanabe,C.K. and

Wood,W.I.  
Secreted and transmembrane polypeptides and nucleic acids encoding the same  
Patent: WO 0116318-A 153 08-MAR-2001;  
Genentech, Inc. (US)  
Location/Qualifiers  
1..1152  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 342 a 244 c 228 g 338 t  
ORIGIN

Query Match 99.8%; Score 1114; DB 6; Length 1152;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tcgagtagaattgtgcgaatggccgcgcctgcagaaatctgtgagctctttcttatgg 60

Db 38 TCAGTTAGAAATGTCTGCAATGGCGCCCTGCAGAAATCTGTGAGCTCTTTCCCTTATGG 97

Qy 61 ggaacctggccaccagctcctctctctgtgcccctctgtgtacagggagagcagctg 120

Db 98 GGACCTGGCCACCAGCTGCCCTCTCTTGGCCCTCTTGTACAGGAGGAGCAGCTG 157

Qy 121 cgccatcagctcccaactgcaggttgacaagtcccaactccagcagccttatcacca 180

Db 158 CGCCATCAGCTCCCACTGCAGGCTTGCAAGTCCAACTTCCAGCAGCCCTATATCACCA 217

Qy 181 acgcacactcagctggtgtaaggagctagcttggctgataaacaacagagcttcgctc 240

Db 218 ACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGCTATAACAACAGAGCTTCGTC 277

Qy 241 tcattggggagaaactgttccacggagtcagtatgagtgagcgtctctctgtatgaagc 300

Db 278 TCATTGGGAGAAACTGTTCCAGGAGTCAGTATGATGAGTGAGCGCTGCTATCTGATGAAGC 337

Qy 301 agtgctgaacttcaaccttgaagaagtgctgttccctcaactgtagtggcttcagcctt 360

Db 338 AGTGCTGAACCTTACCCCTTGAAGAAGTGTCTGTTCCTCAATCTGATAGTTCCAGCCTT 397

Qy 361 atatcagagagtggtgcttccctcctgagcagctcagcaaacaggctaaagcagatgcata 420

Db 398 ATATGAGAGAGTGTGGCTTCCCTTCCCTGGCAGGCTCAGCAACAGGCTAAGCATGTCTATA 457

Qy 421 ttgagagtgatgacctgcataccagaggaatgtcgaagagctgaagcacagagtaaaa 480

Db 458 TTGAAGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTCAAAA 517

Qy 481 agcttgagagagtgagagatcaaaagcaattggagaactggatttgcgttttatgtctc 540

Db 518 AGCTTGGAGAGAGTGGAGAGATCAAAAGCAATTTGGAGAACTGGATTTGCTGTTATGTCTC 577

Qy 541 tgagaaatgcttgcatttgaccagagcaagctgaaataatgaataactaaccccccttcc 600

Db 578 TGAGAAATGCCCTGCATTTGACCAGAGCAAGCTGAAATAATGAATAAATCAACCCCTTTCC 637

Qy 601 ctgctagaataaacaattagatgcccccaagcgatttttttaaccaaagaagatggg 660

Db 638 CTGCTAGAATAACAAATTAGATGCCCAAGCGATTTTTTTTAAACCAAGAAGATGGG 697

Qy 661 aagcacaactccatcatgtggtggtgatttccaaatgaacccctgcgttagttacaagga 720

Db 698 AAGCCAAACTCCATCATGATGGGTGGATTCCAAATTAACCCCTGCGTGTAGTTACAAAGGA 757

Qy 721 aaccaatgcaactttgtttataagaccagaaggtagacttttctaagcagatatttat 780

Db 758 AACCAATGCCACTTTGTTTATTAAGACCAAGAGTACAGCTTTCTAAGCATAGATATTTAT 817

Qy 781 tgataaacttcatgttaactggtgtctctacacagaaacaatttttttaataaa 840

Db 818 TGATAAATTCATGCTTAAGTGGTGTCTATACAGAGAAACAATTTATTTTAAATAA 877



```
Db 1118 TGTTATTGACCTCAATAAACACTGGATATCC 1151
RESULT 5
AX048204
LOCUS AX048204 1132 bp DNA 1 linear PAT 15-DEC-2000
DEFINITION Sequence 44 from Patent WO0070049.
ACCESSION AX048204
VERSION AX048204.1 GI:11876994
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1132)
EXTRACT Azimzai, Y., Yue, H., Lal, P., Burford, N., Bandman, O., Baughn, M. R.,
JOURNAL Extracellular signaling molecules
PATENT: WO 0070049-A 44 23-NOV-2000;
INCYTE Incyte Genomics, Inc. (US)
FEATURES
source
location/Qualifiers
1..1132
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 5571181CB1"
BASE COUNT 333 a 240 c 226 g 333 t
ORIGIN
Query Match 97.7%; Score 1090; DB 6; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1090; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tcgagtagaattgtgcgaatggccctgcagaaatctgtgagctcttcccttatg 60
Db 31 TCGAGTTAGAAATGTCTGCAATGGCCCTGCAGAAATCTGTGAGCTCTTCTTATGG 90
QY 61 ggacctggccaccagctgcctctctctctgtggccctctgtgtacaggaggagcgtg 120
Db 91 GGACCCCTGCCACAGCTGCTCTCTCTTGGCCCTCTTGTGTACAGGAGGAGCAGCTG 150
QY 121 cgcctcatcagctccactcagcgttgacaagtcacactccagcagccttatcacca*180
Db 151 CGCCCATCAGCTCCCACTCAGGCTTGACAAAGTCCAACTCCAGCAGCCTATATCACCA 210
QY 181 accgcaccttcagctggcttaaggagcgtgagctgtgctgatacaacacagagcttcgc 240
Db 211 ACGCACCTTCATGCTGGCTAAGGAGCTAGCTTGGCTGATACACACAGAGCTTCGTC 270
QY 241 tcattgggagaaactgttccacggagtcagtagtgagcgtgctatctgatgaagc 300
Db 271 TCATTGGGAGAAACTGTTCCAGGAGTCAGTATGAGTGCAGCTGCTATCTGATGAAGC 330
QY 301 agtgtcgaacttcaccttgaagaagctgtctccctcaatctgataggttccagcctt 360
Db 331 AGGTGCTGAACCTTACCCCTGGAAGAAGTGTCTCTCCCTCAATCTGATAGGTTCCAGCCTT 390
QY 361 atatgcaggaggtggtgctctctcctgcagcgttcagcaacagcgttaagcacatgcata 420
Db 391 ATATGCAGGAGGTGGTGGCTTCTTCCAGGCTCAGCAACAGGCTAAGCACATGTCATA 450
QY 421 ttgaaggtgatgacctgatccagagaatgtgcnaaagctgaaggacagctgaaaa,480
Db 451 TTGAAGGTGATGACCTTGATATCCAGGAAATGTGCAAGAAAGTGTGAAAAGCTGAAAGCACAGTCAAAA 510
QY 481 agctttgagagagtgagagatacaagcaattggagaactgatttgcgtttattatctc 540
Db 511 AGCTTGAGAGAGTGGAGAGATCAAGCAATTCGAGAACTGGAACTTGTGTTATATCTTC 570
QY 541 tgagaatgcttcgcatttgaccagagcaagcgtgaaaaatgataactaacccctttcc 600
Db 571 TGAGAAATGCCTGCATTTGACCCAGACCAAGCTGAAATGAATACTAACCCCTTCC 630
```

```
QY 601 ctgctgaaataacaattagatgcccccaagcgatttttttaacccaagaagtcgg 660
Db 631 CTGCTAGAAATAACAATTAGATGCCCAAGCGATTTTAAACCAAGGAAGTCGG 690
QY 661 aagccaaactccatcatgagtggtgattccaaatgaacccctgcgttagttacaaagga 720
Db 691 AAGCCAAACTCCATCATGATGAGTGGGATTCCTCAATGAACCCCTGCGTGTAGTTACAAAGGA 750
QY 721 aaccaatgccaactttgtttataagaccagagaagtagagactttctaaagcatagatttat 780
Db 751 AACCAATGCCACTTTTGTTTTATAAGACCAGAGTAGACATTTCTAAGCATAGATATTAT 810
QY 781 tgatacatttcattgtaactggtgtctctatacacagaaacaatttttttaataaa 840
Db 811 TGATAACATTTTCATTTGTAACCTGTTCTATACACAGAAACAATTTATTTTAAATAA 870
QY 841 ttgtcttttccataaaaaagattacttccattccttttaggggaaaaaacccctaaata 900
Db 871 TTGTCTTTTCCATAAAAAAGATTACTTCCATTCCCTTAGGGGAAAAAACCCCTAAATA 930
QY 901 gcttcattgtttccataatcagtcactttatttataaattgtattattattattataaga 960
Db 931 GCTTCATGTTTCCATAATCAGTACTTTATATTATAATCTATTATTATTATTATAAGA 990
QY 961 ctgcattttatttatcatcatttataatgatttatttatttatttatttatttatttatt 1020
Db 991 CTGCATTTTATTATATCATTTTATTAAATATGGATTTATTATTATAGAAACATCATTCGATA 1050
QY 1021 ttgctacttgagtgtaaggctaatattgatttatttattgacataattattagactataaca 1080
Db 1051 TTGCTACTTGAGTGTAAAGGCTAATTATGATATTATATGCAATAATTATTATAGAGCTATAACA 1110
QY 1081 tgtttatttg 1090
Db 1111 TGTTATTATG 1120
RESULT 6
AX054620
LOCUS AX054620 1139 bp DNA 1 linear PAT 13-JAN-2001
DEFINITION Sequence 1 from Patent WO0073457.
ACCESSION AX054620
VERSION AX054620.1 GI:12228184
KEYWORDS unidentified.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1139)
AUTHORS Parham, C. L., de waal Malefyt, R. and Marehalli, N. L.
TITLE Mammalian interleukin-10 homologs: IL-d110 and IL-d210
JOURNAL Patent: WO 0073457-A 1 07-DEC-2000;
SCHERING CORPORATION (US)
FEATURES
source
location/Qualifiers
1..1139
/organism="unidentified"
/db_xref="taxon:3264"
/note="surmised Homo sapiens"
59..598
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC21963.1"
/db_xref="GI:12228185"
/translaton="MAALOKSVSSFLMGTLATSCULLALLVGGGAAAPLSSHCRUDK
SNFOOPYITNRTFMFLAKEXSLADNNTDVRIGELKFLPHGVSMSPERCYLMKOVNFTLEE
VLFPOSDFQPMQEVVPEFLARLSNRLSTCHIEGDDLHLQRNVQKLKDTVKKLGESE
IKRAIGEDLFLFMSLRNACI"
mat_peptide 158..595
BASE COUNT 337 a 238 c 227 g 333 t 4 others
ORIGIN
Query Match 83.7%; Score 934; DB 6; Length 1139;
```

Mon Sep 23 09:43:51 2002

Best Local Similarity 99.7%; Pred. No. 0;									
Matches 1084; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
QY	1	tcgagttagaattgctcaatggcgccctgcagaaatctgtgagctcttcccttatgg	60						
Db	39	TCGAGTTAGAAATGCTGCAATGGCCGCCCTCGAGAAATCTGTGAGCTCTTCCCTATGG	98						
QY	61	ggacctggccaccagctgcctctctctcttggccctcttggtacagggagagcagctg	120						
Db	99	GGACCTGGCCACCAGCTGCCCTCTCTTGGCCCTCTTGGTACAGGGAGGAGCAGCTG	158						
QY	121	cgccatcagctccactgcaggttgacaagtcacaacttccagagcctatatcacca	180						
Db	159	CGCCCATCAGTCCCACTGCAGCTTGCACAGTCCCAACTTCCAGCAGCCCTATCACCA	218						
QY	181	accgacccctcagctggctgaagaggtcgttgcctgctgatacaacacagagcttcgc	240						
Db	219	ACCGCACCTTCATGCTGGCTTGAAGAGNTAGCTTGGCTGATACACACAGAGCTTCGTC	278						
QY	241	tcattggggagaactgttccacggagtcagtgatgagtgagcgtctctatgatgaagc	300						
Db	279	TCATTGGGGAGAAATGCTTCCAGGGAGTCAGTATGAGTGAGCGCTGCTATCTGATGAAGC	338						
QY	301	aggtgctgaactcacccttgaagagtgctgtccctcaatctgatgattccagcctt	360						
Db	339	AGGTGCTGAACTTCACTTGAAGAGTGTCTTCCCTCAATCTGATGTTCCAGCCTT	398						
QY	361	atctgacggagggtggtcccttctcggccggctcagcaacaggtcgaagcacatgtcata	420						
Db	399	ATATGACAGAGGTGGTCCCTTCTGCGCAGGCTCAGCAACAGGCTAAGCACATGTCTATA	458						
QY	421	ttgaagtgatgacctgcatactccagagggaatgtgcaaaagctgaagggacacagctgaaa	480						
Db	459	TTGAAGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAANA	518						
QY	481	agcttgagagagtgagagagatcaacaatttgagagctggttctgttcttatgtctc	540						
Db	519	AGCTTGGAGAGAGTGGAGAGATCAAGCAATTTGGAGAACTGGATTTGCTGTTTATGCTC	578						
QY	541	tgaaaaagctcgtcatttgaccagagcaaaagctgaaaaatgaataaactaaaccccttcc	600						
Db	579	TGAGAAATGCCCTGCTATTGACCAGAGCAAGCTGAAAAATCAATAACTAACCCCTTTCC	638						
QY	601	ctgtagaataaacaattagatgcccacaacgcatgtttttttaaaccaaaagaagatggg	660						
Db	639	CTGCTAGAAATAACAATTAGATGCCCAAGCGATTTTTTTTAAACCAAGAGAGATGGG	698						
QY	661	aagcaaaactccatcatgatgggtggattccaatgaacccctgcgttagttacaaga	720						
Db	699	AAGCCAAACTCCATCATGATGGTGGATTCCAATGAACCCCTCGCTTAGTTACAAAGGR	758						
QY	721	aaccaatgccacttttggttataagaccagaggtagactttctaagcatagatatltat	780						
Db	759	AACCAATGCCACTTTGTTTATAAGACCAGAGGTAGACTTCTWAGCATAGATATTAT	818						
QY	781	tgataaacttccatgtaactggtgtctctatacacagaaaacaattttttttaaataa	840						
Db	819	TGATAAATTTTCAATGTAATGTTCTTATACAGAAAAACAATTTATTTTAAATAA	878						
QY	841	ttgtctttttccataaaaaagattacttccctcttaggggaaaaaaccccttaataa	900						
Db	879	TTGTCTTTTCCATAAAAAAGATTACTTTCCATTCTTTAGGGGAAAAAACCCCTTAATA	938						
QY	901	gcttcacgtttccataaactcagtcactttattattataaaatgattttattattataaga	960						
Db	939	GCTTCATGTTTCCATAATCAGTACTTTATATTTAATATGATTTATTATTATTATAAGA	998						
QY	961	ctgcattttatttatcatcttttataatgattgattttattatagaacaacattcgata	1020						
Db	999	CTGCATTTTATTATATCATTTTATTAATATGATTTATTATTATAGAAAACATATTTCGATA	1058						
QY	1021	ttgctacttgagtgaaaggctaattgtattattatgacaataattatagagctataaca	1080						

Db	1059	TTGCTACTTGAAGCTAATATTGATATTGATATTTATGACAATAATTATAGAGCTATAACA	1118						
QY	1081	tgtttat 1087							
Db	1119	TGTTTAT 1125							
RESULT 7									
HSA277247	HSA277247	Homo sapiens mRNA for Interleukin 22 (IL-22 gene).	690 bp	mRNA	linear	PRI 21-DEC-2000			
LOCUS	AJ277247								
DEFINITION	AJ277247								
ACCESSION	AJ277247.1	GI:9968293							
VERSION	IL-22 gene; Interleukin 22.								
KEYWORDS	human.								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	1 (bases 1 to 690)								
AUTHORS	Dumoutier, L., Van Roost, E., Colau, D. and Renauld, J.C.								
TITLE	Human interleukin-10-related T cell-derived inducible factor: molecular cloning and functional characterization as an hepatocyte-stimulating factor								
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (18), 10144-10149 (2000)								
MEDLINE	20420346								
REFERENCE	2 (bases 1 to 690)								
AUTHORS	Renauld, J.C.								
TITLE	Direct Submission								
JOURNAL	Submitted (10-APR-2000) Renauld J.C., UCL 74 59, Ludwig Institute for Cancer Research, Avenue Hippocrate, 74, B-1200 Brussels, BELGIUM								
FEATURES									
source	1..690								
gene	/organism="Homo sapiens"								
CDS	/db_xref="taxon:9606"								
	/cell_type="peripheral blood mononuclear cells"								
	72..611								
	/gene="IL-22"								
	72..611								
	/gene="IL-22"								
	/note="previously called IL-10-related T-cell-derived inducible factor (IL-TIF)"								
	/codon_start=1								
	/product="interleukin 22"								
	/protein_id="CAC06085.1"								
	/db_xref="GI:9968294"								
	/translation="MAALOKSVSFLMGLTATSCLLALLVOGAAAPISSHICRLDK SNFOOPYITNRTFMIAKEASLADNNTDRLIGKLFHGVSMSEPCYLMKOVLPNTLEE VLFQSDRFQPYMOEVVFFLARLSNRISTCHTEGDDLLHIQRNWKLDYTKKLGESGE IKATGEIDLILFMSLRNACI"								
BASE COUNT	182 a 176 c 167 g 165 t								
ORIGIN									
Query Match 57.3%; Score 639; DB 9; Length 690;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	tcgagttagaattgctgcaatggcgccctgcagaaatctgtgagctcttcccttatgg	60						
Db	52	TCGAGTTAGAAATGCTGCAATGGCCGCCCTCGAGAAATCTGTGAGCTCTTCCCTATGG	111						
QY	61	ggacctggccaccagctgcctctctctcttggccctcttggtacagggagagcagctg	120						
Db	112	GGACCTGGCCACCAGCTGCCCTCTCTTGGCCCTCTTGGTACAGGGAGGAGCAGCTG	171						
QY	121	cgccatcagctccactgcaggttgacaagtcacaacttccagagcctatatcacca	180						
Db	172	CGCCCATCAGTCCCACTGCAGCTTGAAGTCCCAACTTCCAGCAGCCCTATCACCA	231						
QY	181	accgacccctcagctggctgaagaggtcgttgcctgctgatacaacacagagcttcgc	240						
Db	232	ACCGCACCTTCATGCTGGCTTGAAGAGTGTCTTGGCTGATACACACAGAGCTTCGTC	291						



/gene="IL22"  
/frequency="0.01"  
/replace="G"  
2820  
/gene="IL22"  
/frequency="0.45"  
/replace="A"  
3204  
/gene="IL22"  
/frequency="0.05"  
/replace="A"  
3366  
/gene="IL22"  
/frequency="0.35"  
/replace="C"  
3723  
/gene="IL22"  
/frequency="0.02"  
/replace="T"  
3807  
/gene="IL22"  
/frequency="0.01"  
/replace="A"  
4156  
/gene="IL22"  
/frequency="0.01"  
/replace="C"  
4234  
4302  
/rpt\_family="MIR3"  
/rpt\_type="dispersed"  
4290  
/gene="IL22"  
/frequency="0.19"  
/replace="C"  
4303  
4364  
/rpt\_family="MIR"  
/rpt\_type="dispersed"  
4497  
/gene="IL22"  
/frequency="0.05"  
/replace="C"  
4561  
/gene="IL22"  
/frequency="0.05"  
/replace="A"  
4615  
/gene="IL22"  
/frequency="0.02"  
/replace="C"  
4723  
/gene="IL22"  
/frequency="0.47"  
/replace="A"  
4829  
/gene="IL22"  
/frequency="0.01"  
/replace="T"  
5064  
5071  
/gene="IL22"  
/frequency="0.01"  
/replace="A"  
5096  
/gene="IL22"  
/frequency="0.02"  
/replace="T"  
5241  
/gene="IL22"  
/frequency="0.01"  
/replace="C"  
5358  
/gene="IL22"  
/frequency="0.04"  
/replace="C"

variation 5382  
/gene="IL22"  
/frequency="0.15"  
/replace="A"  
5389  
5571  
/rpt\_family="L3"  
/rpt\_type="dispersed"  
5442  
variation  
/gene="IL22"  
/frequency="0.01"  
/replace="T"  
5553  
/gene="IL22"  
variation  
Query Match 57.0%; Score 636; DB 9; Length 8393;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 481 agcttgagagagtgaggagatcaaaacaaattggagaaactggatttgcgtgttattgtctc 540  
Db 6683 AGCTTGGAGAGAGTGGAGAGATCAAGCAATTGGAGAACTGGATTGCTGTTTATGTCTC 6742  
Qy 541 tgagaaatgcctgcatttgaccagagcaaaagctgaaaaatgaataaactcaacccctttccc 600  
Db 6743 TGAGAAATGCCCTGCATTTGACCAGAGCAAAAGCTGAAAATGAATAACTAACCCCTTTCC 6802  
Qy 601 ctgctagaaataacaattagatgccccaaagcgatttttttaacccaaaaggaagatggg 660  
Db 6803 CTGCTAGAAATAACAATTAGATGCCCAAGCGATTTTTTTTAAACCAAAAGAGATGGG 6862  
Qy 661 aagccaaactccatcatgatgggtggattccaaatgaaccctcggttagttacaaga 720  
Db 6863 AAGCAAACTCCATCATGATGGTGGATTCCAAATGAACCCCTCGCTTAGTTACAAAGGA 6922  
Qy 721 aaccaatgccactttgtttataagaccagaggtagacttcttaagcatagatatattat 780  
Db 6923 AACCAATGCCACTTTGTTTATTAAGCCAGAGGTAGACTTTCTAAGCATAGATATTAT 6982  
Qy 781 tgataacatttccattgtaactgggtgtctctatcacagaaaacaaattttttttaaataa 840  
Db 6983 TGATAACATTTTCATTGTAACCTGGTGTCTCTATACACAGAAAACAATTTTATTTTAAATAA 7042  
Qy 841 ttgtcttttccataaaaaagattacttcttccattccttaggggaaacccctcaata 900  
Db 7043 TTGCTCTTTTCCATAAAAAAGATTACTTTCCATTCTTTAGGGGAAAAAACCCCTTAATA 7102  
Qy 901 gcttcatttccataatcagtaactttattataataatgattattattattattataaga 960  
Db 7103 GCTTCATGTTTCCATAATCAGTACTTTATTTATTAATGATATTATTTATTTATATAAGA 7162  
Qy 961 ctgcattttatttatcatcttttataatagtgattttatttagaatacatcattcgata 1020  
Db 7163 CTGCATTTTATTTATATCATTTTATTAATATGGAATTTATTTATAGAAAACATCATTCGATA 7222  
Qy 1021 ttgtcatttgagtgtaaggctaaatattgatattttatgacaataattatagagactataca 1080  
Db 7223 TTGCTACTTGAGTGTAAGGCTAAATTTGATATTTATGACAATAATATATAGAGCTATAACA 7282  
Qy 1081 tgtttatttgacctcaataaaccacttgatattccta 1116  
Db 7283 TGTTTTATTGACCTCAATAAACACTTTGGATATCCTA 7318  
RESULT 9  
AC022511/c AC022511 13350 bp DNA linear PRI 11-JUL-2001  
LOCUS Homo sapiens 12 BAC RP11-71J4 (Roswell Park Cancer Institute Human  
DEFINITION BAC Library) complete sequence.  
ACCESSION AC022511  
VERSION AC022511.22 GI:14669924  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

# REFERENCE AUTHORS

1 (bases 1 to 133350)  
 Alzobrooks, S.L., Amarantunga, H.C., Are, J.R., Banks, T., Barbarella, J., Benton, J., Blum, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowler, S., Brilev, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burnell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Day-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emeling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Garbis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Haves, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseghe, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mel, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokwenkwu, S., Ogih, M., Okwono, G., Oragunye, N., Owiedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojuben, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoostari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Zhou, J., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Zuchiatapati, R. and Gibbs, R.

Direct Submission  
 2 (bases 1 to 133350)  
 Worley, K.C.  
 Direct Submission  
 Submitted (05-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 133350)  
 Worley, K.C.  
 Direct Submission  
 Submitted (11-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Jul 11, 2001 this sequence version replaced gi:12656660.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smil and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT

----- Summary Statistics -----  
 Contig length: 150667  
 Phrap values in estimate: 149705  
 Average error rate (BCM-Phrap estimate): 1.52849e-05  
 Fraction of Phrap values less than 40: 0.0163388  
 Number of consensus changing edits: 83  
 Number of N's in consensus: 0

----- Consensus Changing Edits -----  

Position	Original Context	Edited Context
7824	tattatat(n)nggtatact	tattatat(g)tggtatact
7825	attatat(n)nggtatact	attatat(g)tggtatact
7826	ttatat(n)nggtatact	ttatat(g)tggtatact
7827	ttatat(n)nggtatact	ttatat(g)tggtatact
7828	ttatat(n)nggtatact	ttatat(g)tggtatact
7829	ttatat(n)nggtatact	ttatat(g)tggtatact
7830	ttatat(n)nggtatact	ttatat(g)tggtatact
7831	ttatat(n)nggtatact	ttatat(g)tggtatact
7832	ttatat(n)nggtatact	ttatat(g)tggtatact
7833	ttatat(n)nggtatact	ttatat(g)tggtatact
7834	ttatat(n)nggtatact	ttatat(g)tggtatact
7835	ttatat(n)nggtatact	ttatat(g)tggtatact
7836	ttatat(n)nggtatact	ttatat(g)tggtatact
7837	ttatat(n)nggtatact	ttatat(g)tggtatact
7838	ttatat(n)nggtatact	ttatat(g)tggtatact
7839	ttatat(n)nggtatact	ttatat(g)tggtatact
7840	ttatat(n)nggtatact	ttatat(g)tggtatact
7841	ttatat(n)nggtatact	ttatat(g)tggtatact
7842	ttatat(n)nggtatact	ttatat(g)tggtatact
7843	ttatat(n)nggtatact	ttatat(g)tggtatact
7844	ttatat(n)nggtatact	ttatat(g)tggtatact
7845	ttatat(n)nggtatact	ttatat(g)tggtatact
7846	ttatat(n)nggtatact	ttatat(g)tggtatact
7847	ttatat(n)nggtatact	ttatat(g)tggtatact
7848	ttatat(n)nggtatact	ttatat(g)tggtatact
7849	ttatat(n)nggtatact	ttatat(g)tggtatact
7850	ttatat(n)nggtatact	ttatat(g)tggtatact
7851	ttatat(n)nggtatact	ttatat(g)tggtatact
7852	ttatat(n)nggtatact	ttatat(g)tggtatact
7853	ttatat(n)nggtatact	ttatat(g)tggtatact
7854	ttatat(n)nggtatact	ttatat(g)tggtatact
7855	ttatat(n)nggtatact	ttatat(g)tggtatact
7856	ttatat(n)nggtatact	ttatat(g)tggtatact
7857	ttatat(n)nggtatact	ttatat(g)tggtatact
7858	ttatat(n)nggtatact	ttatat(g)tggtatact
7859	ttatat(n)nggtatact	ttatat(g)tggtatact
7860	ttatat(n)nggtatact	ttatat(g)tggtatact
7861	ttatat(n)nggtatact	ttatat(g)tggtatact
7862	ttatat(n)nggtatact	ttatat(g)tggtatact
7863	ttatat(n)nggtatact	ttatat(g)tggtatact
7864	ttatat(n)nggtatact	ttatat(g)tggtatact
7865	ttatat(n)nggtatact	ttatat(g)tggtatact
7866	ttatat(n)nggtatact	ttatat(g)tggtatact
7867	ttatat(n)nggtatact	ttatat(g)tggtatact
7868	ttatat(n)nggtatact	ttatat(g)tggtatact
7869	ttatat(n)nggtatact	ttatat(g)tggtatact
7870	ttatat(n)nggtatact	ttatat(g)tggtatact
7871	ttatat(n)nggtatact	ttatat(g)tggtatact
7872	ttatat(n)nggtatact	ttatat(g)tggtatact
7873	ttatat(n)nggtatact	ttatat(g)tggtatact
7874	ttatat(n)nggtatact	ttatat(g)tggtatact
7875	ttatat(n)nggtatact	ttatat(g)tggtatact
7876	ttatat(n)nggtatact	ttatat(g)tggtatact
7877	ttatat(n)nggtatact	ttatat(g)tggtatact
7878	ttatat(n)nggtatact	ttatat(g)tggtatact
7879	ttatat(n)nggtatact	ttatat(g)tggtatact
7880	ttatat(n)nggtatact	ttatat(g)tggtatact
7881	ttatat(n)nggtatact	ttatat(g)tggtatact
7882	ttatat(n)nggtatact	ttatat(g)tggtatact
7883	ttatat(n)nggtatact	ttatat(g)tggtatact
7884	ttatat(n)nggtatact	ttatat(g)tggtatact
7885	ttatat(n)nggtatact	ttatat(g)tggtatact
7886	ttatat(n)nggtatact	ttatat(g)tggtatact
7887	ttatat(n)nggtatact	ttatat(g)tggtatact
7888	ttatat(n)nggtatact	ttatat(g)tggtatact
7889	ttatat(n)nggtatact	ttatat(g)tggtatact
7890	ttatat(n)nggtatact	ttatat(g)tggtatact
7891	ttatat(n)nggtatact	ttatat(g)tggtatact
7892	ttatat(n)nggtatact	ttatat(g)tggtatact
7893	ttatat(n)nggtatact	ttatat(g)tggtatact
7894	ttatat(n)nggtatact	ttatat(g)tggtatact
7895	ttatat(n)nggtatact	ttatat(g)tggtatact
7896	ttatat(n)nggtatact	ttatat(g)tggtatact
7897	ttatat(n)nggtatact	ttatat(g)tggtatact
7898	ttatat(n)nggtatact	ttatat(g)tggtatact
7899	ttatat(n)nggtatact	ttatat(g)tggtatact
7900	ttatat(n)nggtatact	ttatat(g)tggtatact
7901	ttatat(n)nggtatact	ttatat(g)tggtatact
7902	ttatat(n)nggtatact	ttatat(g)tggtatact
7903	ttatat(n)nggtatact	ttatat(g)tggtatact
7904	ttatat(n)nggtatact	ttatat(g)tggtatact
7905	ttatat(n)nggtatact	ttatat(g)tggtatact
7906	ttatat(n)nggtatact	ttatat(g)tggtatact
7907	ttatat(n)nggtatact	ttatat(g)tggtatact
7908	ttatat(n)nggtatact	ttatat(g)tggtatact
7909	ttatat(n)nggtatact	ttatat(g)tggtatact
7910	ttatat(n)nggtatact	ttatat(g)tggtatact
7911	ttatat(n)nggtatact	ttatat(g)tggtatact
7912	ttatat(n)nggtatact	ttatat(g)tggtatact
7913	ttatat(n)nggtatact	ttatat(g)tggtatact
7914	ttatat(n)nggtatact	ttatat(g)tggtatact
7915	ttatat(n)nggtatact	ttatat(g)tggtatact
7916	ttatat(n)nggtatact	ttatat(g)tggtatact
7917	ttatat(n)nggtatact	ttatat(g)tggtatact
7918	ttatat(n)nggtatact	ttatat(g)tggtatact
7919	ttatat(n)nggtatact	ttatat(g)tggtatact
7920	ttatat(n)nggtatact	ttatat(g)tggtatact
7921	ttatat(n)nggtatact	ttatat(g)tggtatact
7922	ttatat(n)nggtatact	ttatat(g)tggtatact
7923	ttatat(n)nggtatact	ttatat(g)tggtatact
7924	ttatat(n)nggtatact	ttatat(g)tggtatact
7925	ttatat(n)nggtatact	ttatat(g)tggtatact
7926	ttatat(n)nggtatact	ttatat(g)tggtatact
7927	ttatat(n)nggtatact	ttatat(g)tggtatact
7928	ttatat(n)nggtatact	ttatat(g)tggtatact
7929	ttatat(n)nggtatact	ttatat(g)tggtatact
7930	ttatat(n)nggtatact	ttatat(g)tggtatact
7931	ttatat(n)nggtatact	ttatat(g)tggtatact
7932	ttatat(n)nggtatact	ttatat(g)tggtatact
7933	ttatat(n)nggtatact	ttatat(g)tggtatact
7934	ttatat(n)nggtatact	ttatat(g)tggtatact
7935	ttatat(n)nggtatact	ttatat(g)tggtatact
7936	ttatat(n)nggtatact	ttatat(g)tggtatact
7937	ttatat(n)nggtatact	ttatat(g)tggtatact
7938	ttatat(n)nggtatact	ttatat(g)tggtatact
7939	ttatat(n)nggtatact	ttatat(g)tggtatact
7940	ttatat(n)nggtatact	ttatat(g)tggtatact
7941	ttatat(n)nggtatact	ttatat(g)tggtatact
7942	ttatat(n)nggtatact	ttatat(g)tggtatact
7943	ttatat(n)nggtatact	ttatat(g)tggtatact
7944	ttatat(n)nggtatact	ttatat(g)tggtatact
7945	ttatat(n)nggtatact	ttatat(g)tggtatact
7946	ttatat(n)nggtatact	ttatat(g)tggtatact
7947	ttatat(n)nggtatact	ttatat(g)tggtatact
7948	ttatat(n)nggtatact	ttatat(g)tggtatact
7949	ttatat(n)nggtatact	ttatat(g)tggtatact
7950	ttatat(n)nggtatact	ttatat(g)tggtatact
7951	ttatat(n)nggtatact	ttatat(g)tggtatact
7952	ttatat(n)nggtatact	ttatat(g)tggtatact
7953	ttatat(n)nggtatact	ttatat(g)tggtatact
7954	ttatat(n)nggtatact	ttatat(g)tggtatact
7955	ttatat(n)nggtatact	ttatat(g)tggtatact
7956	ttatat(n)nggtatact	ttatat(g)tggtatact
7957	ttatat(n)nggtatact	ttatat(g)tggtatact
7958	ttatat(n)nggtatact	ttatat(g)tggtatact
7959	ttatat(n)nggtatact	ttatat(g)tggtatact
7960	ttatat(n)nggtatact	ttatat(g)tggtatact
7961	ttatat(n)nggtatact	ttatat(g)tggtatact
7962	ttatat(n)nggtatact	ttatat(g)tggtatact
7963	ttatat(n)nggtatact	ttatat(g)tggtatact
7964	ttatat(n)nggtatact	ttatat(g)tggtatact
7965	ttatat(n)nggtatact	ttatat(g)tggtatact
7966	ttatat(n)nggtatact	ttatat(g)tggtatact
7967	ttatat(n)nggtatact	ttatat(g)tggtatact
7968	ttatat(n)nggtatact	ttatat(g)tggtatact
7969	ttatat(n)nggtatact	ttatat(g)tggtatact
7970	ttatat(n)nggtatact	ttatat(g)tggtatact
7971	ttatat(n)nggtatact	ttatat(g)tggtatact
7972	ttatat(n)nggtatact	ttatat(g)tggtatact
7973	ttatat(n)nggtatact	ttatat(g)tggtatact
7974	ttatat(n)nggtatact	ttatat(g)tggtatact
7975	ttatat(n)nggtatact	ttatat(g)tggtatact
7976	ttatat(n)nggtatact	ttatat(g)tggtatact
7977	ttatat(n)nggtatact	ttatat(g)tggtatact
7978	ttatat(n)nggtatact	ttatat(g)tggtatact
7979	ttatat(n)nggtatact	ttatat(g)tggtatact
7980	ttatat(n)nggtatact	ttatat(g)tggtatact
7981	ttatat(n)nggtatact	ttatat(g)tggtatact
7982	ttatat(n)nggtatact	ttatat(g)tggtatact
7983	ttatat(n)nggtatact	ttatat(g)tggtatact
7984	ttatat(n)nggtatact	ttatat(g)tggtatact
7985	ttatat(n)nggtatact	ttatat(g)tggtatact
7986	ttatat(n)nggtatact	ttatat(g)tggtatact
7987	ttatat(n)nggtatact	ttatat(g)tggtatact
7988	ttatat(n)nggtatact	ttatat(g)tggtatact
7989	ttatat(n)nggtatact	ttatat(g)tggtatact
7990	ttatat(n)nggtatact	ttatat(g)tggtatact
7991	ttatat(n)nggtatact	ttatat(g)tggtatact
7992	ttatat(n)nggtatact	ttatat(g)tggtatact
7993	ttatat(n)nggtatact	ttatat(g)tggtatact
7994	ttatat(n)nggtatact	ttatat(g)tggtatact
7995	ttatat(n)nggtatact	ttatat(g)tggtatact
7996	ttatat(n)nggtatact	ttatat(g)tggtatact
7997	ttatat(n)nggtatact	ttatat(g)tggtatact
7998	ttatat(n)nggtatact	ttatat(g)tggtatact
7999	ttatat(n)nggtatact	ttatat(g)tggtatact
8000	ttatat(n)nggtatact	ttatat(g)tggtatact
8001	ttatat(n)nggtatact	ttatat(g)tggtatact
8002	ttatat(n)nggtatact	ttatat(g)tggtatact
8003	ttatat(n)nggtatact	ttatat(g)tggtatact
8004	ttatat(n)nggtatact	ttatat(g)tggtatact
8005	ttatat(n)nggtatact	ttatat(g)tggtatact
8006	ttatat(n)nggtatact	ttatat(g)tggtatact
8007	ttatat(n)nggtatact	ttatat(g)tggtatact
8008	ttatat(n)nggtatact	ttatat(g)tggtatact
8009	ttatat(n)nggtatact	ttatat(g)tggtatact
8010	ttatat(n)nggtatact	ttatat(g)tggtatact
8011	ttatat(n)nggtatact	ttatat(g)tggtatact
8012	ttatat(n)nggtatact	ttatat(g)tggtatact
8013	ttatat(n)nggtatact	ttatat(g)tggtatact
8014	ttatat(n)nggtatact	ttatat(g)tggtatact
8015	ttatat(n)nggtatact	ttatat(g)tggtatact
8016	ttatat(n)nggtatact	ttatat(g)tggtatact
8017	ttatat(n)nggtatact	ttatat(g)tggtatact
8018	ttatat(n)nggtatact	ttatat(g)tggtatact
8019	ttatat(n)nggtatact	ttatat(g)tggtatact
8020	ttatat(n)nggtatact	ttatat(g)tggtatact
8021	ttatat(n)nggtatact	ttatat(g)tggtatact
8022	ttatat(n)nggtatact	ttatat(g)tggtatact
8023	ttatat(n)nggtatact	ttatat(g)tggtatact
8024	ttatat(n)nggtatact	ttatat(g)tggtatact
8025	ttatat(n)nggtatact	ttatat(g)tggtatact
8026	ttatat(n)nggtatact	ttatat(g)tggtatact
8027	ttatat(n)nggtatact	ttatat(g)tggtatact
8028	ttatat(n)nggtatact	ttatat(g)tggtatact
8029	ttatat(n)nggtatact	ttatat(g)tggtatact
8030	ttatat(n)nggtatact	ttatat(g)tggtatact
8031	ttatat(n)nggtatact	ttatat(g)tggtatact
8032	ttatat(n)nggtatact	ttatat(g)tggtatact
8033	ttatat(n)nggtatact	ttatat(g)tggtatact
8034	ttatat(n)nggtatact	ttatat(g)tggtatact
8035	ttatat(n)nggtatact	ttatat(g)tggtatact
8036	ttatat(n)nggtatact	ttatat(g)tggtatact
8037	ttatat(n)nggtatact	ttatat(g)tggtatact
8038	ttatat(n)nggtatact	ttatat(g)tggtatact
8039	ttatat(n)nggtatact	ttatat(g)tggtatact
8040	ttatat(n)nggtatact	ttatat(g)tggtatact
8041	ttatat(n)nggtatact	ttatat(g)tggtatact
8042	ttatat(n)nggtatact	ttatat(g)tggtatact
8043	ttatat(n)nggtatact	ttatat(g)tggtatact
8044	ttatat(n)nggtatact	ttatat(g)tggtatact
8045	ttatat(n)nggtatact	ttatat(g)tggtatact
8046	ttatat(n)nggtatact	ttatat(g)tggtatact
8047	ttatat(n)nggtatact	ttatat(g)tggtatact
8048	ttatat(n)nggtatact	ttatat(g)tggtatact
8049	ttatat(n)nggtatact	ttatat(g)tggtatact
8050	ttatat(n)nggtatact	ttatat(g)tggtatact
8051	ttatat(n)nggtatact	ttatat(g)tggtatact
8052	ttatat(n)nggtatact	ttatat(g)tggtatact
8053	ttatat(n)nggtatact	ttatat(g)tggtatact
8054	ttatat(n)nggtatact	ttatat(g)tggtatact
8055	ttatat(n)nggtatact	ttatat(g)tggtatact
8056	ttatat(n)nggtatact	ttatat(g)tggtatact
8057	ttatat(n)nggtatact	ttatat(g)tggtatact
8058	ttatat(n)nggtatact	ttatat(g)tggtatact
8059	ttatat(n)nggtatact	ttatat(g)tggtatact
8060	ttatat(n)nggtatact	ttatat(g)tggtatact
8061	ttatat(n)nggtatact	ttatat(g)tggtatact
8062	ttatat(n)nggtatact	ttatat(g)tggtatact
8063	ttatat(n)nggtatact	ttatat(g)tggtatact
8064	ttatat(n)nggtatact	ttatat(g)tggtatact
8065	ttatat(n)nggtatact	ttatat(g)tggtatact
8066	ttatat(n)nggtatact	ttatat(g)tggtatact
8067	ttatat(n)nggtatact	ttatat(g)tggtatact
8068		



Db	1933	1933	TGATAACATTTTCATTGTGAACCTGGTGTCTTACACAGAAAAAATTTATTTTAAATAA	1874
Qy	841	841	ttgtctttttccataaaaaagattactttccattcccttaggggaaaaaacccccaata	900
Db	1873	1873	TTGTCTTTTCCATTAATAAAAGATTACTTTCATTCCTTTAGGGGAAAAAACCCCTAAATA	1814
Qy	901	901	gcttcattgtttccataatccagctacttatattatataatgtattattattattattataa	960
Db	1813	1813	GCTTCATCTGTTTCCATAATTCAGTACTTTATATATTTATAAATGTATTTATTTATTTATATA	1754
Qy	961	961	ctgcattttatttatcatcattttatataataggattttattttatagaaacatcattcgata	1020
Db	1753	1753	CTGCATTTTATTTATATCATTTTATTATATATGGATTTATTTATAGAAACATCATTTGATA	1694
Qy	1021	1021	ttgtacttgagtgtaaggctaatattgatatttatgatataattatagacataatcatagagcata	1080
Db	1693	1693	TTGTACTTGTAGTGTGAAGCTTAATATTTCATATTATATGACAATAATTATAGAGCTATAACA	1634
Qy	1081	1081	tgttatttgacctcaataaaacacttgatataccta	1116
Db	1633	1633	TGTTTATTGTACCTCAATAATAAACACTTGGATATCCTA	1598
RESULT	10			
AC007458/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				



## JOURNAL

Submitted (25-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Jul 31, 1999 this sequence version replaced gi:5230396.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

## COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT-----

```

----- Summary Statistics -----
Contig length: 151624
Phrap values in estimate: 151315
Average error rate (BCM-Phrap estimate): 0.000165522
Fraction of Phrap values less than 40 : 0.020725
Number of consensus changing edits: 5
Number of N's in consensus : 0

```

```

----- Consensus changing edits -----
Position Original*Context Edited*Context
2273 ggccttcatt(n)tcaattgttt ggccttcatt(t)tcaattgttt
5088 agatgggtgc(n)tactctcttc agatgggtgc(c)tactctcttc
102880 gggttcactg(n)gttaactcagg gggttcactg(t)gttaactcagg
123710 tctataatt(n)tggtttatatg tctataatt(t)tggtttatatg
127972 ttttcttgag(n)gaatcctcaag ttttcttgag(t)gaatcctcaag

```

## ----- Distribution of Quality &lt; 40 Bases -----

```

10001
9001
8001
7001
6001
5001
4001
3001
2001
# bases
* * * * *

```

```

1001 * * * * * * * * * *
01 * * * * * * * * * *
5 10 15 20 25 30 35 40
Phrap Value Range

```

## Version: 1.01 qxfo.

## FEATURES

## source

Location/Qualifiers

```

1. 191111
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12q15"
/clone="RPC111-44AB24"
complement(1..3489)
/rpt_family="L1P3"
1..82
/note="Overlaps bases 113681..113762 in clone AC005294"
/function="Overlaps with adjacent clone AC005294"
complement(3702..4041)
/rpt_family="L1P3"
complement(4043..7146)
/rpt_family="L1P8b"
complement(7625..7930)
/rpt_family="L1PA15-16"
complement(8427..8844)
/rpt_family="L1M2"
complement(8860..8988)
/rpt_family="MLTID"
complement(8997..9284)
/rpt_family="L1MC/D"
complement(9927..10097)
/rpt_family="L1M4"
10575..10852
/rpt_family="AluJo"
11022..11060
/rpt_family="AT-rich"
complement(12158..12355)
/rpt_family="MER20"
complement(12407..12592)
/rpt_family="MIR"
12877..12950
/rpt_family="AT-rich"
14057..14169
/rpt_family="MIR"
14920..14949
/rpt_family="(TG)n"
complement(15199..15978)
/rpt_family="L1MA9"
16201..16508
/rpt_family="AluX"
16931..17096
/rpt_family="MIR"
complement(17098..17247)
/rpt_family="MIR"
17645..17841
/rpt_family="MER107"
complement(18996..19099)
/rpt_family="MER91B"
complement(21141..21217)
/rpt_family="L2"
21218..21247
/rpt_family="(CA)n"
complement(21248..21454)
/rpt_family="L2"
21547..21661
/rpt_family="LTR9B"
complement(21670..22020)
/rpt_family="L2"
complement(22021..22378)
/rpt_family="THE1B"
complement(22381..23987)

```

repeat\_region

misc\_feature

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

```

repeat_region      /rpt_family="THEIB-internal"
complement(23989..24349)
/rpt_family="THEIB"
repeat_region      complement(24350..24651)
/rpt_family="L2"
repeat_region      25393..25482

Query Match      57.0%; Score 636; DB 9; Length 191111;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 481 agcttgagagagtgagagatcaaaagcaattggaggaactgattgctgtttatgtctc 540
Db 188243 AGCTTGGAGAGAGTGGAGAGATCAAAAGCAATGGAGAACTGGATTTGCTGTTATGTCTC 188184

Oy 541 tgaagaatgcctgcatcttgaccagagcaaaagctgaaataaactaacccctttcc 600
Db 188183 TCAGAAATGCCCTGCTATTTGACCAAGCAAGCTGAAATGAATCACTAACCCCTTTCC 188124

Oy 601 ctgctagaataaacaattagatgcccaagcgatttttttaacaaaaggaagatggg 660
Db 188123 CTGCTAGAAATTAACAATTAGATGCCCAAGCGATTTTTTTTAAACCAAGGAAGATGGG 188064

Oy 661 aagccaaactccatcatgatgggtggatcccaaatgaaccctcgcttagtagtaacaaagga 720
Db 188063 AAGCCAAACTCCATCATGATGGGTGGATTCCAAAATGAACCCCTGCTGTAGTTACAAAGGA 188004

Oy 721 aaccaatgcactttgtttataagaccagagagtagacttcttaagcatagatttat 780
Db 188003 AACCAATGCCACTTTGTTTATAAGACCAGAGGAGTAGACTTTCTAAGCATAGATATTAT 187944

Oy 781 tgataacatttcattgtaactgggtgtctatcacagaaaaacaatttttttaataaa 840
Db 187943 TGATAACATTTTCATTGTAACGGTGTCTATACACAGAAAAACAATTTATTTTAAATAA 187884

Oy 841 ttgtctttttccataaaaaagattacttctccattccttaggggaaaaaacccctaaata 900
Db 187883 TTGTCTTTTTCATAAAAAGATTACTTTCCATTCTTTAGGGGAAAAAACCCCTAAATA 187824

Oy 901 gcttcattgttccataacagtcactttattattataaattgattattattattataaga 960
Db 187823 GCTTCATGTTTCCATAACAGTACTTTATATTATTAATGATTTTATTATTATTAAGA 187764

Oy 961 ctgcattttatttatcatcttttataatgattgattttattatagaacatcatcgata 1020
Db 187763 CTGCATTTTATTATATCATTTTATTATATGATTTTATTATAGAAACATCATTCGATA 187704

Oy 1021 ttgctacttgagtgtaaggctaatttgatattttatgacaataattatagacctataca 1080
Db 187703 TTGCTACTTGAGTGAAGGCTAATATTGATTTATTATGACAATATTATAGAGCTATAACA 187644

Oy 1081 tgtttatttgacctcaataaaacacttgatatcccta 1116
Db 187643 TGTTTATTGACCTCAATAAACACTTGGATATCCCTA 187608

RESULT 11
AC087562 AC087562 135146 bp DNA linear HTG 19-OCT-2001
LOCUS Pan troglodytes clone RP43-74117, WORKING DRAFT SEQUENCE, 31
DEFINITION unordered pieces.
ACCESSION AC087562
VERSION AC087562.4 GI:14190648
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE chimpanzee.
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE 1 (bases 1 to 135146)
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-olsman, F.R., Allen, C.,
Alabrooks, S.L., Anaratunge, H.C., Are, J.R., Banks, T., Barbara, J.,
Benton, J., Blum, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,

```

```

Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C.,
Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jollivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Louisge, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Mesdor, M.,
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
Sison, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Thomas, N.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Unpublished
2 (bases 1 to 135146)
Direct Submission
Submitted (10-JAN-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza Houston, TX 77030, USA
On May 23, 2001 this sequence version replaced gi:12621390.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: ZUAE
Center clone name: RP43-74117
----- Summary Statistics
Sequencing vector: M13: L08821
Chemistry: Dye-terminator Big Dye: 49% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 134068 bases at least Q40
Consensus quality: 143074 bases at least Q30
Consensus quality: 148054 bases at least Q20
Estimated insert size: 144346; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-gel estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT

\* be preserved.  
\* 1  
\* 10537: contig of 10537 bp in length  
\* 10538 10637: gap of unknown length  
\* 10638 10637: contig of 5859 bp in length  
\* 16497 16596: gap of unknown length  
\* 16597 22635: contig of 6039 bp in length  
\* 22636 22735: gap of unknown length  
\* 22736 28745: contig of 6010 bp in length  
\* 28746 28845: gap of unknown length  
\* 28846 34084: contig of 5239 bp in length  
\* 34085 34184: gap of unknown length  
\* 34185 39046: contig of 4862 bp in length  
\* 39047 39147: gap of unknown length  
\* 39148 45080: contig of 5834 bp in length  
\* 45081 45080: gap of unknown length  
\* 45082 48786: contig of 3606 bp in length  
\* 48787 54719: contig of 5933 bp in length  
\* 54720 54819: gap of unknown length  
\* 54820 59584: contig of 4765 bp in length  
\* 59585 59685: gap of unknown length  
\* 59686 65244: contig of 5560 bp in length  
\* 65245 65344: gap of unknown length  
\* 65345 70092: contig of 4748 bp in length  
\* 70093 70192: gap of unknown length  
\* 70193 73303: contig of 5111 bp in length  
\* 73304 73403: gap of unknown length  
\* 73404 79784: contig of 4381 bp in length  
\* 79785 79884: gap of unknown length  
\* 79885 84754: contig of 4870 bp in length  
\* 84755 84854: gap of unknown length  
\* 84855 88649: contig of 3795 bp in length  
\* 88650 88749: gap of unknown length  
\* 88750 92541: contig of 3792 bp in length  
\* 92542 92641: gap of unknown length  
\* 92642 96461: contig of 3820 bp in length  
\* 96462 96561: gap of unknown length  
\* 96562 100892: contig of 4331 bp in length  
\* 100893 100992: gap of unknown length  
\* 100994 103932: contig of 2940 bp in length  
\* 103933 104033: gap of unknown length  
\* 104034 107186: contig of 3154 bp in length  
\* 107187 107286: gap of unknown length  
\* 107287 110219: contig of 2933 bp in length  
\* 110220 110319: gap of unknown length  
\* 110320 113001: contig of 2682 bp in length  
\* 113002 113101: gap of unknown length  
\* 113102 115998: contig of 2897 bp in length  
\* 115999 116098: gap of unknown length  
\* 116099 119333: contig of 3235 bp in length  
\* 119334 119434: gap of unknown length  
\* 119435 122281: contig of 2848 bp in length  
\* 122282 122381: gap of unknown length  
\* 122382 125109: contig of 2728 bp in length  
\* 125110 125209: gap of unknown length  
\* 125210 127912: contig of 2703 bp in length  
\* 127913 128012: gap of unknown length  
\* 128013 130629: contig of 2617 bp in length  
\* 130630 130729: gap of unknown length  
\* 130730 132831: contig of 2102 bp in length  
\* 132832 132931: gap of unknown length  
\* 132932 135146: contig of 2215 bp in length.  
FEATURES  
source  
1. .135146  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-74117"  
BASE COUNT 40919 a 26449 c 25229 g 39519 t 3030 others  
ORIGIN  
Query Match 22.4%; Score 250; DB 2; Length 135146;  
Best Local Similarity 99.7%; Pred. No. 1.6e-133;

Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 707 ttagttaacaaagaaacaaatccactttgtttataagaccagaggtagactttctaa 766  
|||||  
Db 29041 TTAGTTACAAAGAAACCAATGCCACTTTGTTTATAGACCAGAGGTAGACTTTTAA 29100  
|||||  
Oy 767 gcatagatatttattgataacatttcattgtaactgggtgtctatcacagaaacaatt 826  
|||||  
Db 29101 GCATAGATATTTATTGTATAACATTTTCATTGTAACTGGTCTCTATACAGAAACAAT 29160  
|||||  
Oy 827 tatttttaataaattgctttttccataaaaaaagattacttccattcctttaggga 886  
|||||  
Db 29161 TATTTTAAATAATGCTTTTCCATAAAAAGATTACTTTCCATTCCTTTAGGGGAA 29220  
|||||  
Oy 887 aaacccctaaatagcttcattgttccataaatcagtaactttattataaaatgattta 946  
|||||  
Db 29221 AAAACCCCTAAATAGCTTCATGTTCCATAATCAGTACTTTATTTATAAATGATTTTA 29280  
|||||  
Oy 947 ttattattataagaactgcaattttattattatcatttttataatggattttataga 1006  
|||||  
Db 29281 TTATTATTATAAGACTGCTATTTATTATTATCATCTTTTATTAAATATGATTATAGA 29340  
|||||  
Oy 1007 a 1007  
Db 29341 A 29341  
RESULT 12  
HSA277248 HSA277248 5397 bp DNA linear PRI 30-MAY-2001  
LOCUS Homo sapiens IL-22 gene for interleukin 22, exons 1a-5.  
DEFINITION  
ACCESSION AJ277248  
VERSION AJ277248.1 GI:11967707  
KEYWORDS IL-22 gene; Interleukin 22.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 5397)  
AUTHORS Dumoutier L., Van Roost E., Amey G., Michaux L. and Renaud J.C.  
TITLE IL-TIF/IL-22: genomic organization and mapping of the human and mouse genes  
JOURNAL Genes and immunity. 1 (8), 488-494 (2000)  
MEDLINE 21069354  
PUBMED 11197690  
REFERENCE 2 (bases 1 to 5397)  
AUTHORS Renaud J.C.  
TITLE Direct Submission  
JOURNAL Submitted (10-APR-2000) Renaud J.C., UCL 74.59, Ludwig Institute for Cancer Research, Avenue Hippocrate, 74, B-1200 Brussels, BELGIUM  
FEATURES  
Location/Qualifiers  
1. .5397  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/map="12q15"  
/cell\_line="CESS"  
/cell\_type="EBV-transformed B cell"  
455. .5397  
/gene="IL-22"  
455. .458  
/gene="IL-22"  
485. .489  
/gene="IL-22"  
join(515. .537,626. .857,1289. .1354,1471. .1614,2539. .2604, 5241. .5397)  
/gene="IL-22"  
515. .537  
/gene="IL-22"  
/note="number 1a"  
538. .625  
/gene="IL-22"  
gene  
CAAT\_signal  
TATA\_signal  
mRNA  
exon  
intron

```

exon      /note="number 1a"
626. .857
/ gene="IL-22"
/ note="number 1b"
join(672. 857.1289. 1354.1471. 1614.2539. 2604.5241. .5318)
/ gene="IL-22"
/ note="previously called IL-10-related T-cell-derived
inducible factor (IL-TIF)"
/ codon_start=1
/ product="interleukin 22"
/ protein_id="CAC19409.1"
/ db_xref="GI:11967708"
/ db_xref="SPTREMBL:O9GZX6"
/ translation="MAALQKSVSSFLMGTLATSCLLLLALLVOGGAAAPISHSRLDK
SNFQOQYINRFNFMFLAKEASLADNNTDVLRIKGLFHVGSMSERCYLKQVNLFTLEE
VLFQSDRFPQYMOEVVPLARLSNRLSTCHTEGDDLHQRVQKLKDVVKLGESGE
IKAIGLDLFLWSLRNACI"
858. .1288
/ gene="IL-22"
/ note="number 1b"
1289. .1354
/ gene="IL-22"
/ number=2
/ number=2
1355. .1470
/ gene="IL-22"
/ number=2
1471. .1614
/ gene="IL-22"
/ number=3
1615. .2538
/ gene="IL-22"
/ number=3
2539. .2604
/ gene="IL-22"
/ number=4
2605. .5240
/ gene="IL-22"
/ number=4
5241. .5397
/ gene="IL-22"
/ number=5
BASE COUNT 1525 a 1007 c 1187 g 1678 t
ORIGIN

Query Match 18.5% Score 207; DB 9; Length 5397;
Best Local Similarity 100.0%; Pred. No. 1.6e-108;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tcgagttagaattgtctgcaatggccgcctgcagaaatctgtgagctcttcttattgg 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 652 TCGAGTTAGAAATTCCTGCAATGGCCGCCCTGCAGAAATCTGTGAGCTCTTTCCCTTATGG 711
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 61 ggaccctggcacacagctgctctcttcttggccctcttggtacaggagagcagctg 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 712 GGACCTGGCCACCAAGCTGCCTCTCTCTTTGGCCCTCTTGGTACAGGAGGAGCAGCTG 771
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 121 cgcacatcagctccacagctgacagctgacagctcgaactccagcagcctatatcacc 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 772 CGCCCATCAGCTCCCACTGCAAGCTTGACAAGTCCAACCTTCCAGCAGGCCCTATATCACC 831
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 181 accgcaccttcattgctgctgaagg 207
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 832 ACCGCACCTTCATGCTGGCTAAGGAGG 858
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
AC087562/c AC087562 135146 bp DNA linear HTG 19-OCT-2001
LOCUS Pan troglodytes clone RP43-74117, WORKING DRAFT SEQUENCE, 31
DEFINITION unordered pieces.
AC087562 AC087562
VERSION AC087562.4 GI:14190648
```

## KEYWORDS

SOURCE

ORGANISM

## REFERENCE

AUTHORS

HTG; HTGS\_PHASE1; HTGS\_DRAFT.

Chimpanzee.

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1 (bases 1 to 135146)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,M., Brown,N.P., Bryant,N.P., Buhaey,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Dreher,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 135146)

Worley,K.C.

Direct Submission

Submitted (10-JAN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 23, 2001 this sequence version replaced gi:12621390.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: ZUAE

Center clone name: RP43-74117

----- Summary Statistics

Sequencing vector: M13; L08821

Chemistry: Dye-terminator Big Dye; 49% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 134068 bases at least Q40

Consensus quality: 143074 bases at least Q30

Estimated insert size: 144346; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 31 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 10537: contig of 10537 bp in length  
\* 10538 10537: gap of unknown length  
\* 10638 16496: contig of 5859 bp in length  
\* 16497 16596: gap of unknown length  
\* 16597 22635: contig of 6039 bp in length  
\* 22636 22735: gap of unknown length  
\* 22736 28745: contig of 6010 bp in length  
\* 28746 28845: gap of unknown length  
\* 28846 34084: contig of 5239 bp in length  
\* 34085 34184: gap of unknown length  
\* 34185 39046: contig of 4862 bp in length  
\* 39047 39146: gap of unknown length  
\* 39147 44980: contig of 5834 bp in length  
\* 44981 45080: gap of unknown length  
\* 45081 48686: contig of 3606 bp in length  
\* 48687 48786: gap of unknown length  
\* 48787 54719: contig of 5933 bp in length  
\* 54720 54819: gap of unknown length  
\* 54820 59884: contig of 4765 bp in length  
\* 59885 59884: gap of unknown length  
\* 59886 65244: contig of 5560 bp in length  
\* 65245 65344: gap of unknown length  
\* 65345 70092: contig of 4748 bp in length  
\* 70093 70192: gap of unknown length  
\* 70193 75303: contig of 5111 bp in length  
\* 75304 75403: gap of unknown length  
\* 75404 79784: contig of 4381 bp in length  
\* 79785 79884: gap of unknown length  
\* 79885 84754: contig of 4870 bp in length  
\* 84755 84854: gap of unknown length  
\* 84855 88499: contig of 3795 bp in length  
\* 88500 88749: gap of unknown length  
\* 88750 92541: contig of 3792 bp in length  
\* 92542 92641: gap of unknown length  
\* 92642 96461: contig of 3820 bp in length  
\* 96462 96561: gap of unknown length  
\* 96562 100892: contig of 4331 bp in length  
\* 100893 100992: gap of unknown length  
\* 100994 103932: contig of 2940 bp in length  
\* 103933 104032: gap of unknown length  
\* 104033 107186: contig of 3154 bp in length  
\* 107187 107286: gap of unknown length  
\* 107287 110219: contig of 2933 bp in length  
\* 110220 110319: gap of unknown length  
\* 110320 113001: contig of 2682 bp in length  
\* 113002 113101: gap of unknown length  
\* 113102 115998: contig of 2897 bp in length  
\* 115999 116098: gap of unknown length  
\* 116099 119333: contig of 3235 bp in length  
\* 119334 119433: gap of unknown length  
\* 122281 122381: contig of 2848 bp in length  
\* 122382 125109: contig of 2728 bp in length  
\* 125110 125209: gap of unknown length  
\* 125210 127912: contig of 2703 bp in length  
\* 127913 128012: gap of unknown length  
\* 128013 130629: contig of 2617 bp in length  
\* 130630 130729: gap of unknown length  
\* 130730 132831: contig of 2102 bp in length  
\* 132832 132931: gap of unknown length  
\* 132932 135146: contig of 2215 bp in length.  
\* Location/Qualifiers

FEATURES

source  
1..135146  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-74117"  
BASE COUNT 40919 a 26449 c 25229 g 39519 t 3030 others  
ORIGIN  
Query Match 5.9%; Score 66; DB 2; Length 135146;  
Best Local Similarity 100.0%; Pred. No. 2.6e-26;  
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 417 catattgaagtgatgacctgcataccagaggaatgtgcacacacagt 476  
|||||  
DB 47800 CATATTGAAGTGCATGACCTGCATATCCAGAGGATGTGCAAAAGCTGAAGGACACAGTG 47741  
QY 477 aaaaag 482  
|||||  
DB 47740 AAAAG 47735  
RESULT 14  
AX179614 LOCUS  
AX179614 Sequence 37 from Patent WO0146422. linear PAT 06-AUG-2001  
DEFINITION  
ACCESSION AX179614  
VERSION AX179614.1 GI:15132040  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 778)  
AUTHORS Presnell, S. R. and Kindsvogel, W.  
TITLE Cytokine zcytol8  
JOURNAL Patent: WO 0146422-A 37 28-JUN-2001;  
ZymoGenetics, Inc. (US)  
FEATURES  
source  
1..778  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
47..586  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC49976.1"  
/db\_xref="GI:15132041"  
/translation="MAVLQKSMFSLSMGLTAAASCLLLIALWAQENALPVNTRCKLEV  
SNFQPYIVNRTFMIAKEASLADNNTDVLIGELKLFGRVNAKQDCQYLMKQVLFNLTLE  
VLLPQSDRFQPYMQEVVPELTKLNLQSLSCHISGDDQNTQKNVRLKETVKKLGESGE  
IKAIQELDLLEMSLRNACV"  
BASE COUNT 213 a 193 c 197 g 175 t  
ORIGIN  
Query Match 3.1%; Score 35; DB 6; Length 778;  
Best Local Similarity 100.0%; Pred. No. 2.5e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 471 acagtgaaaaagcttgagagagtgagagagatcaaa 505  
|||||  
DB 497 ACAGTGAAGCTTGGAGAGACTGGAGAGATCAA 531  
RESULT 15  
AR165226 LOCUS  
AR165226 Sequence 7 from patent US 6274710. linear PAT 17-OCT-2001  
DEFINITION  
ACCESSION AR165226  
VERSION AR165226.1 GI:16238719  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 1119)  
AUTHORS Dumoutier, L., Louhed, J. and Renaud, J.  
TITLE Antibodies which specifically bind T Cell inducible factors (TIFs)  
JOURNAL Patent: US 6274710-A 7 14-AUG-2001;

FEATURES  
source Location/Qualifiers  
1..1119

BASE COUNT 352 a 233 c 232 g 302 t  
ORIGIN /organism="unknown"

Query Match 3.1%; Score 35; DB 6; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 2.5e-08;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 471 acagtgaagcttgagagagtgagagatcaa 505  
|||||  
Db 502 ACACTGAAGAGCTTGGAGAGAGTGGAGAGATCAA 536

Search completed: September 23, 2002, 02:50:49  
Job time: 7530 sec